

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 13:46:45 ; Search time 11231 Seconds
3.981 Million cell updates/sec

Title: US-10-664-775-4

Perfect score: 2279

Sequence: 1 gatacctcctctagtgaag.....ttgtaattctagtgctgat 2279

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 0.5

Searched: 45 seqs, 9809608 residues

Total number of hits satisfying chosen parameters: 90

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : rge4.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		ID	Description
		Match	Length		
C	1 840.902	36.9	183326	1	AC124438
	2 840.902	36.9	287253	1	AC147502
	3 780.401	34.2	230813	1	AC114097
C	4 778.202	34.1	192016	1	AC102426
	5 778.202	34.1	227550	1	AC102426
	6 768.303	33.7	217398	1	AC102519
C	7 761.303	33.4	207687	1	AL824706
	8 757.703	33.2	158685	1	AC102201
	9 754.401	33.1	206312	1	AC102739
C	10 752.402	33.0	343010	1	AC120175
	11 752.302	33.0	236604	1	AC099619
	12 752.002	33.0	177414	1	AC124411
C	13 729.201	32.0	184474	1	AL672047
	14 729.201	32.0	204031	1	AC093354
	15 725.001	31.8	211131	1	AL806512
C	16 719.803	31.6	135750	1	AL928859
	17 719.803	31.6	142529	1	AC115852
	18 715.001	31.4	153772	1	AC129304
C	19 715.001	31.4	271609	1	AC122210
	20 714.803	31.4	247119	1	AC126868
	21 714.803	31.4	259629	1	AC119553
C	22 713.403	31.3	235196	1	AC132786
	23 713.403	31.2	247726	1	AC128004
	24 710.102	31.2	245174	1	AC095157
C	25 707.202	31.0	167223	1	AC113514
	26 706.902	31.0	250188	1	AC110100
	27 706.103	30.9	168169	1	AC129113
C	28 704.103	30.9	198631	1	AL683804
	29 704.103	30.9	267360	1	AC101656
	30 703.903	30.9	128936	1	AC015893
C	31 702.902	30.8	211245	1	AC126499
	32 702.902	30.8	241255	1	AC095214
	33 699.902	30.7	227762	1	AL669976

RESULT 1

AC124438	LOCUS	AC124438	183326 bp	DNA	linear	ROD 13-NOV-2003
AC124438	DEFINITION	Mus musculus BAC clone RP24-260K9 from chromosome 7, complete sequence.				
AC124438	ACCESSION	AC124438				
AC124438	VERSION	GI:23462943				
AC124438	KEYWORDS	HTG.				
AC124438	SOURCE	Mus musculus (house mouse)				
AC124438	ORGANISM	Mus musculus				
AC124438	REFERENCE	1 (bases 1 to 183326)				
AC124438	AUTHORS	McLellan, M., Bielicki, L., Dignan, G. and Doeber, A.				
AC124438	TITLE	The sequence of Mus musculus BAC clone RP24-260K9				
AC124438	JOURNAL	Unpublished (2001)				
AC124438	REFERENCE	2 (bases 1 to 183326)				
AC124438	AUTHORS	Wilson, R.				
AC124438	TITLE	Sequencing of Mus musculus				
AC124438	JOURNAL	Unpublished (2001)				
AC124438	REFERENCE	3 (bases 1 to 183326)				
AC124438	AUTHORS	McPherson, J.D. and Waterston, R.H.				
AC124438	TITLE	Direct Submission				
AC124438	JOURNAL	Submitted (14-JUN-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA				
AC124438	REFERENCE	4 (bases 1 to 183326)				
AC124438	AUTHORS	McPherson, J.D. and Waterston, R.H.				
AC124438	TITLE	Direct Submission				
AC124438	JOURNAL	Submitted (01-SEP-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA				
AC124438	REFERENCE	5 (bases 1 to 183326)				
AC124438	AUTHORS	McPherson, J.D. and Waterston, R.H.				
AC124438	TITLE	Direct Submission				
AC124438	JOURNAL	Submitted (02-OCT-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA				
AC124438	REFERENCE	6 (bases 1 to 183326)				
AC124438	AUTHORS	Wilson, R.				
AC124438	TITLE	Direct Submission				
AC124438	JOURNAL	Submitted (13-NOV-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA				
AC124438	COMMENT	On Oct 2, 2002 this sequence version replaced gi:2539265.				

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu>
Contact: submissions@wustl.edu
----- Summary Statistics

Center project name: M_BB0260K09

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:

all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-24 BAC Library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES

```

source
1..183326
   /organism="Mus musculus"
   /mol_type="genomic DNA"
   /db_xref="taxon:10090"
   /chromosome="7"
   /map="7"
   /clone="RP24-260K9"
   /clone_lib="RPCI-24"
repeat_region
421..527
   /rpt_family="L1"
repeat_region
780..849
   /rpt_family="L1"
repeat_region
1229..1489
   /rpt_family="L1"
repeat_region
1925..2147
   /rpt_family="MER1_type"
repeat_region
2269..2345
   /rpt_family="B4"
repeat_region
2432..2639
   /rpt_family="L1"
repeat_region
2657..2747
   /rpt_family="L1"
repeat_region
2753..2819
   /rpt_family="ID"
repeat_region
2825..2891
   /rpt_family="L1"
repeat_region
3225..4557
   /rpt_family="L1"
repeat_region
5567..5680
   /rpt_family="MIR"
repeat_region
6490..6682
   /rpt_family="B2"
repeat_region
6777..6881
   /rpt_family="Alu"
repeat_region
6889..7045
   /rpt_family="B4"
repeat_region
9156..9335
   /rpt_family="MER2_type"
repeat_region
10113..10156
   /rpt_family="B4"
repeat_region
10440..10701
   /rpt_family="B4"
repeat_region
11049..11114
   /rpt_family="ID"
repeat_region
11205..11581
   /rpt_family="L1"
repeat_region
12639..12861
   /rpt_family="B4"
repeat_region
12916..12988
   /rpt_family="B4"
repeat_region
13314..13598

```

```

repeat_region
14985..15253
   /rpt_family="B4"
repeat_region
15284..15355
   /rpt_family="MIR"
repeat_region
15669..15726
   /rpt_family="ERV1"
repeat_region
16139..16229
   /rpt_family="L1"
repeat_region
16418..16525
   /rpt_family="L1"
repeat_region
16880..16932
   /rpt_family="MIR"
repeat_region
16942..17139
   /rpt_family="L1"
repeat_region
17302..17445
   /rpt_family="Alu"
repeat_region
17459..17508
   /rpt_family="L1"
repeat_region
17667..17787
   /rpt_family="Alu"
repeat_region
18127..18260
   /rpt_family="L2"
repeat_region
18386..18472
   /rpt_family="MIR"
repeat_region
18473..18818
   /rpt_family="MaLR"
repeat_region
18831..19229
   /rpt_family="MaLR"
repeat_region
19435..19530
   /rpt_family="MIR"
repeat_region
20110..20262
   /rpt_family="B4"
repeat_region
20395..20447
   /rpt_family="ERV1"
repeat_region
20458..20581
   /rpt_family="MaLR"
repeat_region
21632..21854
   /rpt_family="MER1_type"
repeat_region
21927..22288
   /rpt_family="MaLR"
repeat_region
23000..23102
   /rpt_family="MIR"
repeat_region
23289..23480
   /rpt_family="B2"
repeat_region
23481..23670
   /rpt_family="B2"
repeat_region
23989..24034
   /rpt_family="L2"
repeat_region
24461..24730
   /rpt_family="B4"
repeat_region
25760..25970
   /rpt_family="MER1_type"
repeat_region
26406..26524
   /rpt_family="Alu"
repeat_region
26748..27090
   /rpt_family="MaLR"
repeat_region
28283..28734
   /rpt_family="ERV1"
repeat_region
29104..29319
   /rpt_family="MER2_type"
repeat_region
29316..29425
   /rpt_family="MER2_type"
repeat_region
29597..29747
   /rpt_family="MIR"
repeat_region
29825..29839
   /rpt_family="MIR"
repeat_region
30206..30399
   /rpt_family="MER1_type"
repeat_region
31148..31285
   /rpt_family="MER1_type"
repeat_region
32032..32336
   /rpt_family="MaLR"

```


Query Match 36.98; Score 840.902; DB 1; Length 183326;			
Best Local Similarity 68.68; Pred. No. 0.024;			
Matches 1486; Conservative 0; Mismatches 456; Indels 223; Gaps 26;			
QY	294	TAATTTCTTGATTTCTATCTTTGGCTCAATTTTAACTCAGTAGTAGTGTGTTGGTTCCAT	353
DB	40783	TCCTTTCTAAATTTCTACCTTGACTCAATTTTCAATCTATAATAGATATTTTAGTTCCAT	40842
QY	354	AAGTTTGTAGTTTCTGTTGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	413
DB	40843	GAATTTGTAACTTTC-----TTGTTGTGATATCCAGCTTTAAAGCCATGG	40887
QY	414	TGGTCAGATAGCATAGATATTAATTTCAATTTCTCTTTTATCTGTCGAGACTTGTCTTTG	473
DB	40888	TGGATTAATAGTGTGCGAGGAGTTTATTTCAATTTCTGTTATCTGTTGAGATTTGCTTTG	40947
QY	474	TTTTGAATATGTAATCAATTTTGGAGAGTTTCATAGGGTGTCTGCAAGAAGGTACAGTC	533
DB	40948	TGCTGTAGTATGTCTCAATTTTGGAGAGTTC--CATGAGTGTCTGAGAGGGTATATTT	41005
QY	534	TTTGTGTTTGGTGAAATAGTCTGTAATATCTCTAGTCCACTTGGTTTATGACATCAG	593
DB	41006	TTTTTGTGTCAGTGAAATGTTCTCAAAATTTTAGTCCATTGTTTCTAATGTAAAG	41065
QY	594	TTAGCTCCAGCAATTTCTCTGTTTCTGTTTGTGTTGAGANGACCTPACTGTGGAGAGAT	653
DB	41066	TTAGCTCCAGTAGATACCTGTTTAGCTTTTGTCTGA-ACAACCTATTTATCGGCAAGAT	41124
QY	654	GGGTAATGAGTAGAGCCCACTATCTGTGTGTGAGG-TCAATATGTGATTTTAGCTGTAGC	712
DB	41125	GGAGTATTGAAATCACCCCACTAACTGTGTGAGATCAATATGTGACTTTTAGCTGTAGT	41184
QY	713	TGTGCTCTTTATGAACTTGGGTGACATTTGTTTGGTGATAGACATTAAGAAATGCA	772
DB	41185	AGTGTCTTTTATGAACCTTGAAGACCTGTTTTCGGTGCAGAGATAATAAGAAAGCA	41244
QY	773	ATGTCCTCTTGGTGA-----TTTTCTTTGATGCGCTATGTAGTATTTCTCCCAATCTCATC	829
DB	41245	ATAATCTCTTATATAATTTTTCCTTTGATTAATTAATGTAGTGTCTTTCCCTATCTCTTC	41304
QY	830	TGCTTAGTTTGGGTTTAAAGTCTAT-TAGTCAGATATTAATAAGTACTGTAT--CGGCTTG	886
DB	41305	TGAATAGTTTGAATTTGAAGTCTGTCTCATCAATAGTAATAAGTACTATATACAGCTTG	41364
QY	887	CTTCTTAGGGCAATTTGCTTAGAATATC-TTTTCCATCCTTTTACTCTAAGTGATGTCT	945
DB	41365	CTTGTAGGTCCATTTGCTTCAATATCTTTTCCATCCTTTTACTCTTAGGTAAATGACT	41424
QY	946	ATCC---ATGGTAGTGTCTTTTGGATGCGAGTAGGATGGAT-CTTGTGTTTCATA	1001
DB	41425	ATCCCTTGATGTCAGGTGCAATTTCTTGAATACAGCAGAAGAAATGGATCCCTGTTTTCACA	41484
QY	1002	TCCATTTCTGTTACCCAGTATCTTTTCTAGAGAAATTAAGATCAATGAGTCAATGATGTT	1061
DB	41485	CTTATCTCTGTAGCCTATGTTTCTTACTGGTAACCTAAGA-----CTACTGATCTT	41536
QY	1062	GAGATTAATCAATGAGCAGTGTGTTGGAATCTTGTATCTGTGACTTGTGAAGTGCTG	1121
DB	41537	AAAAAGTGTATGAGCAGTCTCTGTTGATTTCTGCTGCTCTTTTCTGTTGTGTGTGTG	41596
QY	1122	TG	1181
DB	41597	TCTGTGTGTGTTTGTGTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	41656
QY	1182	TGCTGTGTGTCT-----	1192
DB	41657	TGTGTGTGTCTGTGCT	41716
QY	1193	-----	1192
DB	41717	TGCTGTGTCTCTGCT	41776
QY	1193	----TGTGTGTCTGTGTTCTCTCCCTCTTCTTGTATTTTGGCTGGAAATATTTTATTTAT	1248
DB	41777	TGTGTGTGTGTGTGTGTTCTCTCCCTCTTCTTTTAAATTTGCTGATCTGGGATTT-----TTCTT	41831
QY	1249	CATATTTTCTTGAATGTGGTAAACATCTTTAGATTTGAAGTTTCTCTCTAG--CCTTCTT	1306
DB	41832	TGTGTGTTTCTTCTGAGTGTGGTTAAACATCTTTACAGATGAATAATTTCTCTCTAGCACCTTATA	41891
QY	1307	TAGGTCTGCAATTTGAAGATAGATATTTCTTTTCACTGATTTTTTATCTTTAGAAATGTCTTCT	1366
DB	41892	TAGGCTGAATTTGTATATAGATATTTGTTAAATTTGATTTTATCTTAGAA---TAGCT	41947
QY	1367	TTCTCCAACTATTTGTGACAGAAAGTTTTTCTPAAGTGCAGTAGTCTGGCTGACATCTGTA	1426
DB	41948	TTCTCCATCTTTTGTAAATAGTAAATTTTCTAAGTACAGTAGCTGGCTGGCTGACATCTGTG	42007
QY	1427	GTCTCTTGAGTCTGTGACACATCTGTGCAGGGCTTCTTTACATTTTGTAGTTTCTATTGG	1486
DB	42008	GGCTCTTAGAGTCTGTGGATCAATTTGTCCAGGCCCTTCTGGCTTTTAGAGTCTCCAAT--G	42066
QY	1487	AAAGTCAAGTGTAAATTTCTPAATACATCTGCCCTTTATATGTTAAATTTGTTTCTTCTT	1546
DB	42067	TGAATTCAGATGTTCTGTCTATAGGGTCTGCCCTTTGTATATTACTTTGCCCTTTTCCCTTG	42126
QY	1547	GCATCTTTTAAATATCTTTTCTTGTCTTATCTTTTAGTGATTTGATTTATGCACTGT	1606
DB	42127	CAGCTTTTAGAATCTTTCTTTTCTGTACTTTTCTTTTAGTGTTTGGATTTAGTCTCTCA	42186
QY	1607	GGGAGCTTTC-TTTTCCGGTCCAACTATTTGGTGTGTTTGTATGCTTCTGTACCTGTAT	1665
DB	42187	GAGATTTTCTTTTCTAGTCTCATCTCTTTAGAGTCTGTATGCTTCTTTAGTCTGTAT	42246
QY	1666	AGGATCTCTTCTCAAGTGTAGGAATTTTCTTTTGTGTTTGTGTTTCTTCTGAAATATTTTC	1725
DB	42247	AGGATCTCTTCTCT-----GTTAGGGAATTTTCTTCTCTGACTTTGTTGAAATATTTTC	42302
QY	1726	CTGCTTTTGAAGCTG-CCTTCTTCCCTTCTCTCTATTC-----CTTTGGTGT	1771
DB	42303	TGTGCTCTGACTGTGAGTGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	42362
QY	1772	TTGCATAGTGTCTCTGGCTTCTGTGATGTTTATGCTCTGA-----	1812
DB	42363	TTGCAAGTGTCTCCAGGTTTCTCTGGATGTTTATGCTGGATTTGGGGTGGGGTGTCTGT	42422
QY	1813	-----TTATTTTAGACTTAACATTTTCTTTGACCAAGGTATCCATTTCTCTCTCTCTCTCT	1867
DB	42423	TTGTTTGTGTTTGTATTAACATTTCTTTGACTAAAGTATCCATTTCTCTCTCTCTCTCTCT	42482
QY	1868	TTCACTGCTGAGATTTCT	1927
DB	42483	TTCACTCTGAGAGA---GTCT	42538
QY	1928	GGTTCCTG-TTGGGCTCTTAATTTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1986
DB	42539	GGTTCCTGATGTAGTGTCT	42598
QY	1987	TATTAATCTCTATTTCCACTTTCCAGTCTCTGAAATGTTTACTCATTT---TTCTCCCACT	2043
DB	42599	TATTAATCTCTATTTCCCACTTTCCAGGCTTTGAAATGTTTATTTTCTCTCTCTCTCTCTCT	42658
QY	2044	ATTATCAATTTTCATGATTTTCTTTTAAAGATTTTATTTCTCTCTCTCTCTCTCTCTCTCTCT	2103
DB	42659	GTITGTATTTTCTAGGTTTCTCTTAATAGATTTTATTTTATTTTCTCTCTCTCTCTCTCTCT	42718
QY	2104	TGAATTCATAAAATGATGTTAAAGTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT	2163
DB	42719	AATATTTTAAAGAGCTATGTTAAAGTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT	42778
QY	2164	GGGCTTATGTAA-----TAGGGTTTTCAGGAGACATATGCTCTGCTCTGCTCTGCTCTGCT	2214
DB	42779	GGGCTTATGTAAATGATGTTGTTAGGGCTCTAGTGGTGGACATAGTGTCTCTTTTGAATATTA	42838
QY	2215	GTCTGTGTTTTGCTTTGGCATATAGACGGCTGAGTTGGGATGATTTGTAATTTCTAGGTG	2274

[illegible]

Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorensueta, L., Louisedge, H., Lozada, R.J., Lu, X., Ma, J.,
 Maheshwari, M., Mainardine, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E.,
 Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
 Narkervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K.,
 Nwankwelen, O., Okunolu, G., Olampunsgoon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C.,
 Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
 Puzo, M., Quirroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
 Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajic, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Stearns, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willison, R., Wleczek, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstock, G., and Gibbs, R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 230813)
 Worley, K.C.

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

Submitted (07-MAR-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 230813)

REFERENCE

AUTHORS

JOURNAL

Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

COMMENT

On May 13, 2003 this sequence version replaced gi:24819108.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequencing contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GLIA

Center clone name: CH230-90E1

----- Summary Statistics

Assembly program: Atlas 3.0:

Consensus quality: 226424 bases at least Q40

Consensus quality: 227854 bases at least Q30

Consensus quality: 228759 bases at least Q20

Estimated insert size: 240103; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 1 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 230813: contig of 230813 bp in length.

FEATURES

source

Location/Qualifiers
 1. 230813
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-90E1"

Query Match 34.2%; Score 780.401; DB 1; Length 230813;

Best Local Similarity 68.5%; Pred. No. 0.061;

Matches 1509; Conservative 0; Mismatches 511; Indels 184; Gaps 33;

QY 207 TTACCATGCTCTCTCTCTCCCTTCTCTAAACACTTCTGGCCAGGCTAGGGCACTACC 266

Db 116565 TTGCTTCAAACTGCTCTTATTGCTTTTCCATGACCTTTGGCATGTTGATTTTCATTTTC 116506

QY 267 GCATTCCTCTCTCTTCCAAACACTTCTATTCTTCTGATTTCTATCTTTGCTCA-TTTTTA 325

Db 116505 ATTTGATTTTAGAAGTTTTTTTAAACCTTCTTCTTAATTTCTGCTTGACTCATTTTTT 116446

QY 326 ACTCAGTAGTGAGTGTGTTGGTTTCCATAAGTTGTAAGTTTCTGTTCTTTCTGTTGTT 385

Db 116445 ATTTAGTGGTGAGTGTGTTTAGTTTTCATGAATTT-----TTTGTAACTTTCT 116398

QY 386 GTTGTGTTTATCTATAGATTTTAAAGCTGTGTGGTGCAGATAGGATAGATATTATTTCAT 445

Db 116397 GTTGTGATATCCAGCTTTTATCCATGGTGGTATATATAGGTGCAGGAATATTTCAT 116338

QY 446 TGTCTTTATCTGTCAGACCTTGTCTTTGTTTGAATAATGATTTCAATTTTGGAGATTT 505

Db 116337 TTTCTTTGATCCACTCAGATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 116278

QY 506 CATAGGTGTGTGACAAAGGTACAGTCTTGTGTTTGGTCAAAATAGTCTGTAATATC 565

Db 116277 TCCATGCACTGATGAGAAAGTATATTTCTTTGAGTTCAAGTGAATATTTCTGTAATAG 116218

QY 566 TTAGTGTCACTTGGTTTATGACATCAGTAGTCTCAGCAATTTCTGTTTCTGTTTGTG 625

Db 116217 T-TAGTGTCAATTTAGTTTATAAGTAGTAGTCTCAGTATTTCTGTTTGTAGC---TTTTG 116162

QY 626 TTGAGATGACCTAACCTGTTGGAGAGATGGGTATTGAAGTAGCCACATCTGTGTGTG 685

Db 116161 TCTGAAGCACTGTTTATTGGCAAGATAGAGTACTGAAATCACCACATACACTGTGTG 116102

QY 686 AGG-TCAAATATGTGATTTTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 744

Db 116101 AGGATCAATATGTGACTTTAGCTTTAGTAATGTTTCTTTATGAACCTGGGTGACCTGT 116042

QY 745 GTTGTGTGATAGATTAAGAAATGCAATGTCCTTCTGTGTGA---TTTTTCTTTGATG 801

Db 116041 ATTTCAGTGCATAGATTAAGAAATGCAATGTCCTTCTGTGTGAATAATTTTCTTTGTGA 115982

QY 802 CCTATGTAGTATTTCTCCCAATCTCATCTGTCTAGTTTGGTTTGGTTTAAAGTCTAT-TAGTCA 860

Db 115981 TGTACGTAGTGTCTTCCCT 115922

QY 861 GATATTAATGACTGT--ATCGGCTTGTCTTCTTAGGGCCATTTCTCTAGATATCTTTT 918

Db 115921 GATAGTAAATGGCTATACACAGCTTTGTTGTAGGTCTATTTGTTCTCAATATCTTTT 115862

QY 919 CCATCTCTTTTACTCTAAGGTGATGCTATCC---ATGGTAGGTGTTCTTTTGTGATGCA 975

Db 115861 CCATCTTTTACTCTTAGATATGACTATCTCTTGGTTTAAAGTCTAT-TAGTCA 115802

QY 976 GCAGTAGGATGAT-CTTGTGTTTCTATCCATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1034


```

Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., MacLean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
McEltrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramaasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Wu,X.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wyman,D.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (31-JUL-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 192016)
Birren,B., Nusbaum,C. and Lander,E.
Direct Submission
Submitted (19-AUG-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
5 (bases 1 to 192016)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukgafter,B., Camarata,J., Chang,J., Choepel,Y.,
Collamore,A., Cook,A., Cooke,P., Corum,B., DeAtellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,B., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kellis,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Liu,X., Mabbitt,R., MacLean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
McEltrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramaasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wyman,D.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (27-SEP-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 19, 2003 this sequence version replaced gi:33354396.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: Li8775
Center clone name: 88_A_10
-----
FEATURES
Source
1..192016
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosomes="1"
/map="1"
/clone="RP24-88A10"
/misc_feature 1..28181
/clone_lib="RP24 Male Mouse BAC"
/note="wgs end extension"
clone_end:SP6"
3..3738
/rpt_family="RLTRETN_MM-int"
complement(1432..1436)
unsure
/note="<30 qual SINGL region"
complement(1454..1817)
unsure

```

```

unsure
/note="single clone coverage"
complement(1464..1510)
/note="<30 qual single clone coverage"
complement(1609..1618)
unsure
/note="<30 qual single clone coverage"
complement(1652..1656)
unsure
/note="<30 qual single clone coverage"
complement(1942..1947)
unsure
/note="<30 qual SINGL region"
complement(1968..1975)
unsure
/note="<30 qual SINGL region"
complement(1976..2096)
unsure
/note="single clone coverage"
complement(1976..1981)
unsure
/note="<30 qual single clone coverage"
complement(2427..2597)
unsure
/note="single clone coverage"
complement(2598..2603)
unsure
/note="<30 qual SINGL region"
complement(2702..2714)
unsure
/note="<30 qual SINGL region"
complement(2723..2731)
unsure
/note="<30 qual SINGL region"
complement(2737..2763)
unsure
/note="<30 qual SINGL region"
complement(2771..2779)
unsure
/note="<30 qual SINGL region"
complement(2780..2805)
unsure
/note="<30 qual single clone coverage"
3739..4057
/rpt_family="RLTRETN_MM"
4058..4195
/rpt_family="B1_MM"
5186..5333
/rpt_family="B1_MM"
complement(5378..5777)
/rpt_family="RLTR9C"
5818..5940
/rpt_family="L1"
5942..5969
/rpt_family="(TTAA)n"
complement(5970..6055)
/rpt_family="RSINE1"
6797..6824
/rpt_family="(TTTG)n"
complement(6826..6979)
/rpt_family="B1_MM"
7587..7606
/rpt_family="(CAA)n"
complement(7608..12044)
/rpt_family="Lx6"
12045..12079
/rpt_family="(TTG)n"
complement(12080..12138)
/rpt_family="Lx6"
12139..12195
/rpt_family="(TG)n"
complement(12196..13147)
/rpt_family="Lx6"
16881..16789
/rpt_family="RSINE1"
16790..16832
/rpt_family="(CA)n"
17775..17796
/rpt_family="AT rich"
complement(18890..19727)
/rpt_family="Lx8"
19864..19886
/rpt_family="AT rich"
complement(20072..20486)
/rpt_family="BGLII"
20608..20829
/rpt_family="ORR1D"

```

Qy	1178	TCTGTGTCGTGCTGTGTGTGTGTGTGTCTCTCCCTCTTTTGGATTTTGGCTCGAAT	1237
Db	12177	TGTGTGTGTGTG---TGTGTGTCTATTCCCTCTCTTTGATT---TTTCTGTCTCGGAT	12230
Qy	1238	TATTATTATTCAATATTTCTTGAATGTGGGTAAACATCTTTAGATGTGAAGTTTTCTCTCT	1297
Db	12231	TATTCATTTTCTGTATTATCTGTGAATCGAGTGAACCTCTTTAGATTGGTCTTTTCCCTCT	12290
Qy	1298	AG--CCTCTTTAGTCTGCATTTGGAAGATATATCTTTACATCTG-ATTTTATCTTA	1354
Db	12291	AGAATCTCTGTAGGTTGGATTTATAGATAGATATGCTTTAAATTTGTTTTTATCATG	12350
Qy	1355	GAATGCTTTCTTTCTCCAACTATTGTGCACAGAAAGTTTTCTAAGTGCAGTAGCTGGC	1414
Db	12351	GGATGCTTTGTTTTCTCCATCTATGGTGTATAAAAAGTTTTCCAAAGTATAGTCTTGG	12410
Qy	1415	CTGACATCTGTAGTCTCTTGGAGTCTGTAGACATCTGTGAGGCGCTTCTTACATTTG	1474
Db	12411	ATAGCATCTGAGTCTCTTAGAATCTGCAACACATCTCTCAGGTCTCT--CTGGCTTTTA	12468
Qy	1475	AGTTTCTATTGGAAAGTCAGGTGTAAATCTTAATACATCTGCTTTATATGTTAAATGGT	1534
Db	12469	GATCTCCATTTGAGAGCTCAGGTGTGTCTTACGAGGTCTGACITTAATGTTACTTGGT	12528
Qy	1535	CTTTTTTCCCTGCATCTTTTAAATCTTTCTTTGTCTTAPACTTTTGTAGTATTTGATT	1594
Db	12529	GTT-----ACTTGTAACTTTTAATACTTTTCTTTGTCTCTGTATGTTAGTGTTTAAT	12583
Qy	1595	ATTATGCATGTGGGAGTTCTTTTCCGGTCCAAATCATTTGGTGTTTCTATGCTTCT	1654
Db	12584	ATTATGTCCAAAGGCTTTCTTTTCTAGTCCAGGCTGTTTGGTGTCTGTATGGTCT	12643
Qy	1655	TGPACTTTGATAGGCATCTTTCTCAAGGTTAGGAAATTTTCTTTTTTGGTTTTCTTG	1714
Db	12644	TGTTCTTTTATATACATCTCTTCCATAGGTAGGACA--TTTCTTCCATAACTTTGTGT	12702
Qy	1715	AAATATTTTCCCTGTTTTGACCT-GCCTCTTCCCTCTCTCATTTCTTCTTGTT---	1770
Db	12703	AAAAGATTTTCCANGCTTTGATCTGGTTCTCTTCTTCTCTATTTCTTATGATCTT	12762
Qy	1771	-----TTTGATAGTGTCTCTGGCTTCTCGATGTTTTATGCTCGATTATTTTA	1820
Db	12763	AGGTTTGACCTTTACATAGTATCTCAAATTTCCAGGATGTTTGTGTGAGAACTGTTTTA	12822
Qy	1821	GACTTAACATTTCTTTGACAGGATATCCATTTCTCTATCTGTGCTTCACTGCCCTGAG	1880
Db	12823	GATTTAGCATTTCTTTTGACTGATGTGCCATTTCTCTATCATATCTTTCTGCGCTCAG	12882
Qy	1881	ATTCTCTCTTATCTTTGTATTTCTGTCT-AGTGAAGCTTGTCTCTGAGGTCTCTG-ITG	1938
Db	12883	ATTCTATCTTCCACCTCTGTATTCATCTTTTGAAGATGTCTTTGTGGTTCCTGTTTG	12942
Qy	1939	GGTTCTTTAATTTTTCATTTCCAGATTTCTCTCAGTTTGGGTTTTGTTTATTAATCTAT	1998
Db	12943	AGTCTCTAAGTTTTTATTCTTAGCTTAAGCTTGGTTTTGGGTTTTCTTTATTTGATCTAT	13002
Qy	1999	TTCCATTTTCAGGTCTGAAATGTTTTTACTCATTTTTCTCCAGTATTTACATTTTCAATA	2058
Db	13003	TTCCATTTTCAGGTCTTGAAACAATTTA--TTAGTTTTCTTTCACGTTTTGTGTTTTCATG	13061
Qy	2059	GGTTCTTTAATGATTTATTCATTTCTCTTCAAGGACCTTTTATGAATTTCAATAAATG	2118
Db	13062	GATTTCTTTAAAGATTTTATTTCTTTTAAAGATCTCTATCATATCCATACAGGC	13121
Qy	2119	TATGTTAAGTCTCTGCCCTTGTCTTACGATAGTT-GCATTTCTCAGGCGCTATTGTAAT	2177
Db	13122	TTTTTTTAAATATTTTCTTTGGATTCAGGTATGTGGAAATGCTTGTGGCTGTGCTAGT	13181
Qy	2178	A-----GGGTTTTAGAGGGAATATGTCTCGTTGTATTGTCTGTGTTTTTCTCT	2229
Db	13182	AGAATTGCTTGAATCTTAGTGGAGACACATGTCTCGGCTGTATTGTGTTTTTTTTTTT	13241

TITLE
JOURNAL
REFERENCE
AUTHORS

Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zalcoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 217398)

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArelhiano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fato, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Harez, N., Hagopian, D., Hegos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., MacLean, C., MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Melarini, J., Menes, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zalcoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (09-JUN-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 9, 2003 this sequence version replaced gi:22380823.
All repeats were identified using RepeatMasker:
Smit, A. F. A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RN/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L18952
Center clone name: 560_1_9
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 206419 bases at least Q40
Consensus quality: 211445 bases at least Q30
Consensus quality: 212661 bases at least Q20
Insert size: 182000; agarose-fp.
Insert size: 213698; sum-of-contigs
Quality coverage: 12.1 in Q20 bases; agarose-fp
Quality coverage: 10.3 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 38 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
* 1 19129: contig of 19129 bp in length
* 19130 19229: gap of 100 bp
* 19230 19836: contig of 607 bp in length
* 19837 19936: gap of 100 bp
* 19937 20613: contig of 677 bp in length
* 20614 20713: gap of 100 bp
* 20714 21355: contig of 642 bp in length

21356 21455: gap of 100 bp
21456 22430: contig of 975 bp in length
22431 22530: gap of 100 bp
22531 23146: contig of 616 bp in length
23147 23246: gap of 100 bp
23247 23990: contig of 744 bp in length
23991 24090: gap of 100 bp
24091 24657: contig of 567 bp in length
24658 24757: gap of 100 bp
24758 25477: contig of 720 bp in length
25478 25577: gap of 100 bp
25578 26032: contig of 515 bp in length
26033 26192: gap of 100 bp
26193 26836: contig of 644 bp in length
26837 26936: gap of 100 bp
26937 27400: contig of 464 bp in length
27401 27500: gap of 100 bp
27501 28196: contig of 696 bp in length
28197 28296: gap of 100 bp
28297 28967: contig of 671 bp in length
28968 29067: gap of 100 bp
29068 29828: contig of 761 bp in length
29829 29928: gap of 100 bp
29929 30975: contig of 1047 bp in length
30976 31075: gap of 100 bp
31076 31970: contig of 895 bp in length
31971 32070: gap of 100 bp
32071 33032: contig of 962 bp in length
33033 33132: gap of 100 bp
33133 33910: contig of 778 bp in length
33911 34010: gap of 100 bp
34011 34616: contig of 606 bp in length
34617 34716: gap of 100 bp
34717 35694: contig of 978 bp in length
35695 35794: gap of 100 bp
35795 36680: contig of 886 bp in length
36681 36780: gap of 100 bp
36781 37711: contig of 931 bp in length
37712 37811: gap of 100 bp
37812 38862: contig of 1051 bp in length
38863 38962: gap of 100 bp
38963 39805: contig of 843 bp in length
39806 39905: gap of 100 bp
39906 41567: contig of 1662 bp in length
41568 41667: gap of 100 bp
41668 43156: contig of 1489 bp in length
43157 43256: gap of 100 bp
43257 44564: contig of 1308 bp in length
44565 44664: gap of 100 bp
44665 46143: contig of 1479 bp in length
46144 46243: gap of 100 bp
46244 47645: contig of 1402 bp in length
47646 47745: gap of 100 bp
47746 49089: contig of 1324 bp in length
49090 49169: gap of 100 bp
49170 51131: contig of 1962 bp in length
51132 51231: gap of 100 bp
51232 52967: contig of 1736 bp in length
52968 53067: gap of 100 bp
53068 56770: contig of 3703 bp in length
56771 81259: contig of 24389 bp in length
81260 81359: gap of 100 bp
81360 114087: contig of 32728 bp in length
114088 114187: gap of 100 bp
114188 158558: contig of 44371 bp in length
158559 158658: gap of 100 bp
158659 217398: contig of 58740 bp in length.

Location/Qualifiers
1. .217398
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"

FEATURES
source

```
/clone="RP24-56019"
/clone_lib="RP24-24 Male Mouse BAC"
1. .19129
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
19230. .19836
/note="assembly_fragment"
19337. .20613
/note="assembly_fragment"
20714. .21355
/note="assembly_fragment"
21456. .22430
/note="assembly_fragment"
22531. .23146
/note="assembly_fragment"
23247. .23990
/note="assembly_fragment"
24091. .24657
/note="assembly_fragment"
24758. .25477
/note="assembly_fragment"
25578. .26092
/note="assembly_fragment"

Query Match 33.7%; Score 768.303; DB 1; Length 217398;
Best Local Similarity 67.6%; Pred. No. 0.068; Indels 75; Gaps 20;
Matches 1349; Conservative 0; Mismatches 572;

QY 292 TCTATTTCTGATTTCTATCTTGGCTCATTTTAACTCAGTAGTCAGTGTGTTGGTTTCC 351
DB 126640 TTCTTTCTTAATTTCTATTTGGCT-AGGTTTAAATCAATAGACTTCTTCAATTTCC 126698

QY 352 ATAAGTTTGAATTTTCTGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 411
DB 126699 ATGAGTTTGAAGTTTCTCAATGATTC--TGTTGTTGTTGTTGTTGTTGTTGTTGTT 126755

QY 412 GGTGTCAGATAGGACATAGATATTTTCAATTTGCTTTTATCTGTCGAGACTTGCTT 471
DB 126756 GGTGTTAGATAGATACAGGGTGATTTTGAATTTCTGTTGTTGTTGTTGTTGTTGTT 126815

QY 472 TGTGTTGAAATGATTTCAATTTGG--AGATTTTCAATAGGGTCTGACAGAGTAC 529
DB 126816 TGTGTTGATGATGTTGGCAATTTTGGAAAGTTCCATGTTGCTGAGAAGAGGCAC 126875

QY 530 AGTCTT-TGTGTTTGGTGAATAGTCTGTAATATCTCTAGGTCCTAGTGGTTATGAC 588
DB 126876 ATCTCTGTTTGTGGAATGTTCTGTAGATATATGTTAGTCCAAATGTTTATAAC 126935

QY 589 ATCAGTTAGCTCCAGCATTTCTCTGTTTCTGTTTGTGAGATCACCTAACTGTTGAG 648
DB 126936 TTCAGTCAGCTCAGCGCTTCTCTGTTAAATTTTGGCTG-GATGGCTGTCCATGGCA 126994

QY 649 AGAATGGGTATGAAATAGCCACTATCTGTGTGTA-GGTCAATATGCTATTTAGCT 707
DB 126995 AGAGTGAATATGCAATTTCCCATTTATAGTTTGTGAGGGTCAGAAATGCTATTAAC 127054

QY 708 GTAGCTGTGTTGTTTATGAACTGGGTGACATTTGTTTGTGTCATAGACATTAAGAA 767
DB 127055 GTAGAAGTTTCTTTATGAACTTGGTACCTTGGTTTGGGCTATAGAT 127114

QY 768 TTGCAATGCTCTTGTGGTATTTTCCCTTTGATGCTATGATGATTTCTTCCCAATCTCA 827
DB 127115 TGGAAATGATGCTGTGGTGAATTTTCCCTTCAATGAATATATAGTATCTTCTGCTAT 127174

QY 828 TCTGCTTATGTTTGGTTTATGCTATTA-CTCAGATATTAATAGTATGCTGATCGGTTG 886
DB 127175 TCTGATAGTTTGGTTTGAAGTCTATTTTGTGAGATTAATTAATGCGTTAACTATCTTT 127234

QY 887 CTCTTAGGCCATTTGCTTGAATATCTTTTCCATCTTTTACTCTAAGGTGATGCTA 946
DB 127235 CTCTTAAGTCTATTTGCTTTGAATATTTTCCAACT--TTACCTGACATTAATGCTA 127292

947 TC---CATGGTAGTCTGCTTTTGGATGACAGCAGTAGGATGATCTGTTTTCATATC 1003
127293 TCTTTGATGTTAAGATGTTTACAGGCACAGCAGAGGGTGAACCTGTTTITAGATC 127352
QY 1004 CATCTGTTACCCAGTATCTTTTCTAGAGAAATTAAGATCATTTAGTCAATGATGTTGA 1063
DB 127353 CATCTGTTAGTCTGTTTCTTTTATTTAGGAATTTGAG-----GCTATTGATATTA 127404
QY 1064 GAATTATCAATGACAGCTGTTTGGGATCTTCTGATCTTATCTGCACTTCTGAGTGTGTTG 1123
DB 127405 GAATAGCAATGGAATTTGTTTCTTACTCTATTTATTTTGTGTAGTAGAAGTAGTAAT 127464
QY 1124 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1183
DB 127465 GTTTGTTGCCATTTGTTGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 127513
QY 1184 TCTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1243
DB 127514 TATGTTTCTTTTCT-----TTTGAATTTCTGGTCTGGGATTTT 127554
QY 1244 TTATTCATATTTTCTTGAATGTTGGTAAACATCTTTAGATTTGAAGTTTCTTCTAG--CC 1301
DB 127555 TTCCCTTTGTTTGTGTGATGTAGTTAAACATCTTTAAGTTGGAGTTTACCTTGTAGCACC 127614
QY 1302 TTCTTTAGGCTGCAATTTGAAGATAGATATCTTTACATCTGATTTTATCTTAAAGTCTC 1361
DB 127615 TTCTAGAGGATGAATTTGTAGATAAATATTTGCTTAAATTTGTTTGTGATGGAATCTT 127674
QY 1362 TTCTTTCTCCAACTATTTGTGACAGAAAGTTTCTTAAAGTGCAGTAGTCTGGCTGACAT 1421
DB 127675 TCATTTATCCATCTGTCGGGTGCGCACTTTGGTTGTATAGTTGTCTGTCAGGCA- 127733
QY 1422 CTGTAGTCTCTGAGTCTGTAGACATCTGTGACGGGCTTCTTACATTTTCAAGTTTCT 1481
DB 127734 -TTTGTCTCTGGAACATGAAGCAGATCTATCTAGTGCCTTTTGGCTTTTGAATCTCC 127792
QY 1482 ATTGGAAGTCAAGTGTAACTTCTATACATCTGTCCTTTATATGTTAAATTTGGTCTTTT 1541
DB 127793 ATT-GAGAAATCAGATATAATCTTAATAGTCTGCTTTATATGTTTGTCTTGTCT- 127850
QY 1542 CCCTTGCACTTTTAAATATTTCTTTTGTCTATATCTTTTGTAGTATTTGATTTATATGC 1601
DB 127851 CCCTGACAGCTTTTAAATATTTCTTTTGTCTATACATTTATTTTCAATTTATATGA 127910
QY 1602 ACTGTGGGAGTTTCTTTCCGCTCAATCTATTTGGTGTGTTTGTATGCTTCTTTACCT 1661
DB 127911 GCCAAGAGGAATTTATTTTGTCTAGTCTTTTGTGTATGTTTGTACTTTTGTACT 127970
QY 1662 TGATAGCATCTCTTTCTCAAGGTTAGGAAATTTTCTTTTGTGTTTCTTCTGAAAATAT 1721
DB 127971 TGATAGCATATCATTTCTTTTACATTTAGGAAATTTTCTTCTATGTTTATTTGAAAAT 128030
QY 1722 TTTCCCTGCTTTTGTACCTGCTTTTCCCTTCTCTATTTCC-----TTTGGTT 1770
DB 128031 TTTTGTACCTTCAATCTGTTTCTTCTCTGATCTTATCTCTATTTATTTATTTAGCC 128090
QY 1771 TTTGTCATAGTCTCTGCTGCTTCTGATGTTTATGTCCTGGATTTATTTAGACTTTAACT 1830
DB 128091 TTTTACAGAGTCAAGTTTCTCGGTGTTTATACAGAGTGTCTTTAGATTTAAAT 128150
QY 1831 TTTTCTTGAACCAAGTATCCATTTCTCTATCTTGTCTTCACTGCTGAGATTTCTCTCT 1890
DB 128151 TTTTCTTGAACCAATATAATTTACTTTCTACCATTTATGTTTCAATTTGCTGAAATTTCTCT 128210
QY 1891 CTATCTCTGTTATTTGTGTCAGTGTGCTGAGGTTCTCTGAGGTTCC---TGTGGTCTTTAA 1947
DB 128211 CCAATGCTTTGTTCTGTGGGTGAAGCTTGTCTGATTTCCATGAGGATTTCTTTA 128270
QY 1948 TTTTCTTCAATTTCCAGATTTCTTCAAGTTTGGGTTTGTGTTTAAATTTATTTTCCACTTT 2007
DB 128271 ATTCTTCACTGACAGATTTCTCTAAGTTTGGGTTATCTTTATTTGATTTCTATCTACT 128330
QY 2008 CAGGTCTGAAATGTTTACTCATTTTCTCCAGATTTTACATTTTCTTCAAGGTTTCTTT 2067
```

Accession	Gene	Strain	Position	Sequence
128331	Db	128389	128389	CAGGCTTGAATGATCTTATTCA-TTTCTTTCCATCAATTTGCTTTCATAAATTTGTT
2068	Qy	2127	2127	AATGAGATTATTCATTCTCTCTTCAAGACACCTTTTATGAATTCATAAAATGATGTTAAG
128390	Db	128449	128449	AAAAAATTTATTTATTTCTCTTTAAGGATCTCCACCATATTTGTAAAAAGCTATTTTAAG
2128	Qy	2187	2187	GTCTCTTGCTTGCTGCTTCAGCTATGTCATTCACGGCCTATTGTAATAGGGTTTTAG
128450	Db	128508	128508	GTTTTTGTCTTAAGTTTCAACCAATGTTGGAATAT-TCAGTGTCAAGTTGCTGGGCTATCA
2188	Qy	2247	2247	CAGGACATATTTGCTCGGTTGTTATTTGCTCTGTGCTTTTTGCTTTGGCATATAGACGGCTG
128509	Db	128568	128568	TGGACACATATTTGCTCTGGTTATTATTATTGATTTGTTGTTGTTTCAATCAATTTGTTTTTG
2248	Qy	2263	2263	AGTTTGGGATGATTCT
128569	Db	128584	128584	GTTTTGGGTTTAGGGT

```

RESULT 7
AL824706
LOCUS
DEFINITION
AL824706 207687 bp DNA linear ROD 17-APR-2003
Mouse DNA sequence from clone RP23-218A4 on chromosome 4, complete
sequence.
ACCESSION
AL824706
VERSION
AL824706.8 GI:30024422
KEYWORDS
HTG.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryote; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 207687)
AUTHORS
Tracey, A.
REFERENCES
Direct Submission
TITLE
Submitted (17-APR-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerv@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Apr 18, 2003 this sequence version replaced gi:25045345.
COMMENT
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerv@sanger.ac.uk

```

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TRMBE; Wp, WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/projects/C.elegans/wormpep> RP23-218A4 is from the RPCI-23 Mouse BAC Library constructed by the group of Pister de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBACE3.6.

FEATURES

Location/Qualifiers

```

1. 207687
/o:ganisms="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="4"
/clone="RP23-218A4"
/clone_lib="RPC1-23"

Query Match      33.4%; Score 761.303; DB 1; Length 207687;
Best Local Similarity 68.6%; Pred. No. 0.071;
Matches 1387; Conservative 0; Mismatches 477; Indels 159; Gaps 26;

300 TTGATTTCTATCTCTGGCTCA---TTTTAACTCAGTAGTAGTGTTGGTTTCCATAAG 356
Db  TCTATTTCTTTCTTGACCATTTTTTTTCATTCAGTAGTGAATTGTTCCATGCTGTAAG 114165

357 TTGTAAGTTTTCGTTGTTCTCTGTTGTTGTTGTTGTTTATCTAGATTTAAGCTGTGGTGG 416
Db  TTTGTAAACTTTCTGAGGTTTCTGCTGGTGGTGATAT--CCCACTTTAAATCCACTGTGG 114223

417 TCAGATAGGACATAGAGTATTATTCAATTGCTTTTATCTGTCGAGACTTGCTTTGTTT 476
Db  TCAGATAGGATGACAGGCGGTATTATTGATTTCTTCTATCTATTGAGACTTGCCTTGTGT 114283

477 TGAATATGTATTCAATTTTTGGAGAGTTTTCATAGGTGCTGCAAGAAGTACAGTCTTT 536
Db  CTGAGCATATGGTGTATTTTGGAGAAAGTTCTCAGGTGCTGAGAAAAAG-----TGT 114336

537 GTGTTTGGTGAATAGTCTGTAATAATCTCTAGGTCCACTTGGTTTATGATCAGATTGA 596
Db  ATGTTTGTGTAAGGTTCTGTTAATA--GGCTAGGTCTTTATGGTTTATAGCCTCAGTTA 114395

597 GCTCCAGCATTTCTCTGTTTCTGTTTTGTGAGATGACCTAACTGTTGGAGAGAAATGGG 656
Db  GCTCCAGCATTTCTCTATTTAGTTTTGTTGCTG--GATGACCTGCTATTGGTGAGAGTGA 114454

657 GTATTGAAGTAGCCCATATCTGTGTGTGAGG--TCAATATGATTTTACGTGTAGCTGT 715
Db  GTATTGAAGTCTCCCATGTCAGTCAGTGTGTGAGGTCGAATGTGATTTAAGCCCTAGCAGT 114514

716 GCTGTTTTATGAACCTTGGGT--GACATTGTGTTTGGTGTGATACACAAATTAAGAATTGCAAT 774
Db  ATTTCTCTATGAACCTTGGGTAACCTTGTTGTTGGTGTGATAGCTCTTGAAGCTTGAAT 114574

775 GPTCTCTTTGGTGA--TTTTCTTTGATGCCATGTATGTATCTTTCCCAATCTCATCTGT 833
Db  GTCTTCTTGGTGTACTTTTTCTTTGATGACTATGCAGTGCTTCCCTCATCTCTCTGAT 114634

834 TAGTTTTGGGTTTAAGTCTATTAG--TCAGATATTAAATGACATGTATCGGCTTGCTTCT 892
Db  TAGTTTTGGTTGAAGTGTGTATTTTAAAAATTAATAAGTGTACACAGCTTGTCTTCT 114694

893 AGGGCCATTTGCTTAGAATATCTTTCCATCCCTTTTACTCTAAGGTGAATGTATATCCATG 952
Db  AGGTCCATTTGCTGTAAATACCAATTT--TCCATCCACCCTGGGTAATGTCTACCCCTTG 114751

953 ---GTAGTGTCTTTTTTGGATCCAGCAGTAGGATGGATCTTTGTTTTCATATCCATCTCT 1009
Db  CTATTAAATGTGTATTTCTTGGATCAACAGAAAGATGGATTCGCGCTTTTACATCCATCTCT 114811

1010 GTTACCACAGTATCTTTTCTAGAGAAATTAAGATCATTTGAGTCAATTCATGTTTGAGAAATTA 1069
Db  ATTAGTCTGTGTCITTTATATGGGAATTGA-----GACCACAGATGTTGAGAGATA 114862

1070 TCAATGACAGAGTTTTGPGAAATCTTGTTATCTTGCACCTGTGGAAGTGTGTGTGTGTG 1129
Db  TCAATGATTAATGATAGTAAATTCCTGTTAT-----TGTAGTGTGTATGTGT--- 114909

1130 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1189
Db  -----TGATATGTGTGTGTATTG 114929

1190 TCTTGTGTGTGTGTCTCTCTCCCTCTTTTGTGATTTTGGCTGGAAATTAATTAATTATTC 1249

```


JOURNAL	Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE	3 (bases 1 to 206312)
AUTHORS	Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., DeAstellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyns, S., Gordon, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, C., MacLean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Minova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zemlek, L., Zimmer, A. and Zody, M.
TITLE	Direct Submission
JOURNAL	Submitted (20-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT	On Aug 20, 2002 this sequence version replaced gi:17060849. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html ----- Genome Center Center: Whitehead Institute/ MIT Center for Genome Research Center code: WIBR Web site: http://www-seq.wi.mit.edu Contact: sequence_submissions@genome.wi.mit.edu ----- Project Information Center project name: L19385 Center clone name: 309_H3 ----- Summary Statistics Sequencing vector: Plasmid; n/a; 100% of reads Chemistry: Dye-terminator Big Dye; 100% of reads Assembly program: Phrap; version 0.960731 Consensus quality: 204656 bases at least Q40 Consensus quality: 205335 bases at least Q30 Consensus quality: 205573 bases at least Q20 Insert size: 187000; agarose-fp Insert size: 205712; sum-of-contigs Quality coverage: 10.3 in Q20 bases; agarose-fp Quality coverage: 9.4 in Q20 bases; sum-of-contigs ----- * NOTE: This is a 'working draft' sequence. It currently consists of 7 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved. * 1 671: contig of 671 bp in length * 672 771: gap of 100 bp * 772 989: contig of 218 bp in length * 989 1039: gap of 100 bp * 1039 18168: contig of 17079 bp in length * 18168 18269: gap of 100 bp * 18269 53423: contig of 35155 bp in length * 53423 53524: gap of 100 bp * 53524 94083: contig of 40560 bp in length * 94083 94183: gap of 100 bp * 94183 142338: contig of 48155 bp in length * 142338 142439: gap of 100 bp * 142439 206312: contig of 63874 bp in length. Location/Qualifiers 1. .206312 /organism="Mus musculus" /mol_type="genomic DNA"

Talimas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,
 Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (23-APR-2003) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Apr 17, 2003 this sequence version replaced gi:28604153.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L25752
 Center clone name: 492_C_11

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 18 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 50033: contig of 50033 bp in length
 * 50034 50133: gap of 100 bp
 * 50134 50435: contig of 302 bp in length
 * 50436 50536: gap of 100 bp
 * 50536 51197: contig of 662 bp in length
 * 51198 51297: gap of 100 bp
 * 51298 52656: contig of 1359 bp in length
 * 52657 52756: gap of 100 bp
 * 52757 57009: contig of 4253 bp in length
 * 57010 57109: gap of 100 bp
 * 57110 60822: contig of 3713 bp in length
 * 60823 60923: gap of 100 bp
 * 60923 66293: contig of 5371 bp in length
 * 66294 66394: gap of 100 bp
 * 66394 72812: contig of 6419 bp in length
 * 72813 72913: gap of 100 bp
 * 72913 82696: contig of 9784 bp in length
 * 82697 82797: gap of 100 bp
 * 82797 94299: contig of 11503 bp in length
 * 94300 94399: gap of 100 bp
 * 94400 118094: contig of 23695 bp in length
 * 118095 14391: contig of 100 bp
 * 14392 14391: contig of 25297 bp in length
 * 14392 14391: gap of 100 bp
 * 14392 167313: contig of 23721 bp in length
 * 167313 167413: gap of 100 bp
 * 167413 180073: contig of 12661 bp in length
 * 180074 180174: gap of 100 bp
 * 180174 198870: contig of 18597 bp in length
 * 198871 198970: gap of 100 bp
 * 198971 261067: contig of 62896 bp in length
 * 261067 302119: gap of 100 bp
 * 302119 302120: contig of 40953 bp in length
 * 302120 302220: gap of 100 bp
 * 302220 343010: contig of 40791 bp in length.

FEATURES

source
 1..343010
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="RP24-492C11"
 /clone_lib="RPCI-24 Male Mouse BAC"
 1..50033
 /note="assembly_fragment
 clone_end:SP6
 vector_side:left"

misc_feature 50134..50435
 /note="assembly_fragment"
 misc_feature 50536..51197
 /note="assembly_fragment"
 misc_feature 51298..52656
 /note="assembly_fragment"
 misc_feature 52757..57009
 /note="assembly_fragment"
 misc_feature 57110..60822
 /note="assembly_fragment"
 misc_feature 60923..66293
 /note="assembly_fragment"
 misc_feature 66394..72812
 /note="assembly_fragment"
 misc_feature 72913..82696
 /note="assembly_fragment"
 misc_feature 82797..94299
 /note="assembly_fragment"
 misc_feature 94400..118094
 /note="assembly_fragment"
 misc_feature 118195..14391
 /note="assembly_fragment"
 misc_feature 143592..167312
 /note="assembly_fragment"
 misc_feature 167413..180073
 /note="assembly_fragment"
 misc_feature 180174..198870
 /note="assembly_fragment"
 misc_feature 198971..261066
 /note="assembly_fragment"
 misc_feature 261167..302119
 /note="assembly_fragment"
 misc_feature 302220..343010
 /note="assembly_fragment"
 Query Match 33.0%; Score 752.402; DB 1; Length 343010;
 Best Local Similarity 70.5%; Pred. No. 0.12;
 Matches 1426; Conservative 0; Mismatches 456; Indels 140; Gaps 33;
 Qy 300 TTGATTTCTATCTTGGCTCA--TTTAACTCAGTAGTACAGTGTGTTGGTTTCCATAAGT 357
 Db 75922 TTGATTTCTTCTTGACCAATTTTTCATTCAGTAGTACAGTGTGTTGGTTTCCATAAGT 75863
 Qy 358 TTGTAAGTTTCTGTTGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 417
 Db 75862 TTGTAAGTTTCTGTTGTT--TTATATTGTTGATACCTCAGCTTTGATCTGTGGTGGT 75807
 Qy 418 CAGATAGGACATAGAGTATTATTTCAATTGTTCTTTTATCTGTCGACACTGCTTTGTTT 477
 Db 75806 CAGATAGTATGCGAGGGACTATTTCATTTCTTGTATCTGTGAAACCTGCTCTGATC 75747
 Qy 478 GAAATATGATTCAATTTTGGAG--AGTTTCATAGGTTGCTGACAAAGAGGTACAGTCTT 535
 Db 75746 TGAATATAGCTCAGTTTGGAGAAAGTCCATAGTACTGAGAGAAAGTATTTCT 75687
 Qy 536 TGTGTTTGTGAAATAGTCTGTAATATCT-CTAGGTCCACTTGGTTTATGACATCAGT 594
 Db 75686 TGTGTTGGGTGAAATATCTGTAATATTTATTAGGTCTATTGTTGTTTATAATGTTCTT 75627
 Qy 595 TAGCTCCAGCATTTCTGTTTCTGTTTGTGAGATGACCTAACTGTTGGAGAAATG 654
 Db 75626 TAACCTCAGTATTCTC-----TTTAGTCTTGTCTGATTAACCTATTATTGAGAGT 75574
 Qy 655 GGGTATTGAAAGTAGCCCACTATCTGTGTGAGGTCAATATGATTATTAGTGTGAGCTG 714
 Db 75573 GAGTATTGAAGCTTCCCACTGTGAGAATGAAG-----TATCAATTAGTAC 75528
 Qy 715 TGCTTGTGTTTATGAACCTTGGGTGACATTGTTTGTGTCATAGACATTAAGATTGCAAT 774
 Db 75527 TGTTCCTTTTTCGAACCTTAGGTTCCCTTGGTTGGTGAATAGATAAT-ACAATTGCAAT 75469
 Qy 775 GTCCTCTTGGTGA--TTTTCTTTTGATGCCCTATGATGATTATTTCCCAATCTCATCTGCT 833

Db 75468 GTCCCTTTGTTGGATTTTCTTTGATGTATATAGTGCCTT---TATCTTTTCTGAT 75412
Qy 834 TAGTTTGGGTTTAAGTCTA-TTAGTCAGATATAAATAGATGATCGGCTTGTCTTT 892
Db 75411 TASTTTTGGTTCAAGTCTAATTTTGTAGATATAAATAGGCTGACAGCTCACTTCA 75352
Qy 893 AGGGCAATTTGCTTGAATATCTTTTC-CATCCTTTTACTCTAAGGTGATGTCATCC-- 949
Db 75351 AGGTCATTTTCTTGGATATCTTTTCTTATCTTCTTTTACCTGAACTGATGCTATTCTT 75292
Qy 950 -ATGCTAGTGTCTTTTCTTGGATGACGACGATGAGATCTTGTCTTTCATATCCATTC 1008
Db 75291 GATGTTAAGTGTCTTCTTGGGTACAGCAAGGATGGATCCTGTTTTCATCCATTC 75232
Qy 1009 TGTATCCAGTATCTTTTCTAGAGAAATTAAGATCAATGATGATGATGATGATGAT 1068
Db 75231 AGTTAGTCTGTCTTTTACTGGGAACTGA-----TACCATTATGTTGAGAT 75181
Qy 1069 ATCAATGACGATGTTTGTGATCTTGTATCTGTCATCTGTCAGATGTCGTCGTGTGT 1128
Db 75180 TTAATGACGATGTTTGTGATCTTGTGATCTTGTATCTTGTGTTGTTGTTGTTGTTGTT 75121
Qy 1129 GTGCTGT 1188
Db 75120 CTCATGTGTACATGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 75061
Qy 1189 GTCTGT 1248
Db 75060 TTTTGATTTGTTGTCT-----GGGATTATTCCT 75032
Qy 1249 CATATTTCTGAAGTGGTAACATCTTTAGATGAGATTTTCTCTCTAG-CCTTCTT 1306
Db 75031 CAAGTTTCTTGAGTTTGTATTAACCTCTTTAGGTGGAGATCTTTCTAGAACTTCTGT 74972
Qy 1307 TAGGCTGATTTGAAGATAGATATCTTTACATCTGATTTTATCTTGAATGCTTTCT 1366
Db 74971 TAGGCTGGATTTGACATAGATATGTTTAAATTTGTTTATCGGGGTTGTTTAT 74912
Qy 1367 TTCTCCACTATTGACAGAAAGTTTCTTAAGTCAGTGTAGTCGGCTGACATCTGTA 1426
Db 74911 TTCCCATCTATCATGACTGAAGTTTCTCGAGTAAAGTGTGAGTCTGGACTGGTA-CTATA 74853
Qy 1427 GTCTCTTGAGTCTCTAGACATCTGTGACGGCTCTCTACATTTTGTGTTTCTATTGG 1486
Db 74852 GTTCTAGAGTCTGACGATATCTG-CCTATGCTCTAGCTTTTAGAGTATCATGG 74794
Qy 1487 AAAAGTCAGGTAAATCTATAFACATCTGCTTATATGTAATGTTGCTTTTTCCTT 1546
Db 74793 -GAAGTCAGGTCTGACGATATCTG-CCTATGCTCTAGCTTTTAGAGTATCATGG 74736
Qy 1547 GCATCTTTTAAATCTTCTTCTTGTCTATATCTTTAGTGTATGATTTATGCACTGT 1606
Db 74735 GCATCTTTTAAATA-----TTCCTGTCTGTATGTTTAGTGTATTTATATGT- 74686
Qy 1607 GGGGAGTTCTTTCCGTCACATCTATTGGTGTGTTTGTATGTTCTTGTACCTTGATA 1666
Db 74685 GGGGACTTCTCTCTGTTCCAGTCTATTGGTGTCTGTAAAGCTCTTGACCTTTATA 74626
Qy 1667 GCATCTCTTCTCAGGTAGGAATTTTCTTTTGTGTTTCTTGTGAAATATTTTCC 1726
Db 74625 GTCACTTTT---TAGGTTAAGCAATTTTCTTCTATGATTTTGTCTGAAATATTTTCT 74569
Qy 1727 CTGCTTTTGAAGCT-GCCTTCTTCCCTCTCTATTC-----TTTGGTTT 1772
Db 74568 GTGCTTTTGAAGCTGGTTTCTTCTCTTATCTATCTATCTATCTTAGATTTGTTG 74509
Qy 1773 TGCATAGTCTCTGCTCTCTGATGTTTATGCTGGATTTATTTAGACTTAACTT 1832
Db 74508 TTTAGTGTCTTCCAGACTTTCTGGATGTTTGTGCGCAAGTTTTATAGAC-----TTT 74455
Qy 1833 TCTTTGACCAAGGATTCATCTTCTATCTTCTTCTTCTGCTGCTGATTTCTCTTCT 1892
Db 74454 TCTTTGATGAGGAT-CATTTCTTCTTATTTGTTCTTCAAGTCCCAAGA-----CTCTTCT 74400

Qy 1893 ATCTCTTGTATCTGTGACGAGGCTTGTCTCTGAGTTCCTCT-TGGGTTCTTAATTTT 1951
Db 74399 GTCTCTTGTATCTGTGATGATGTTTACCTCAGAGTTCCTGTGAGTTCCTAATAT 74340
Qy 1952 TTCAATTTCCAGATTCCTTCCAGTTTGGGTTTCTTTTATTAAATCTATTTCCACTTTTCCAG 2011
Db 74339 TTTGCTTCCCAATTTTCTCAG-TTAGGTTTCTTTTATTGATCTATTCTCTCTTCCAG 74281
Qy 2012 TCTGGAAT-GTTTACTCATTTCTCCAGTATTTACATTTTCAAGTTCCTTTTAAAT 2070
Db 74280 TCTGGAATTTTATTAAATTTCTTCCACTGTTTGTGTTTTCATAGATTTTCCCTAAA 74221
Qy 2071 GGATTTATTCCTTCTTCAAGGACCTTTTATGAATTCATAAAATGATGTTAAAGTC 2130
Db 74220 GGATATATTCATTTCTTAAAGTCTCTGTCACATTCATAAGACGCTTTTAAGTC 74161
Qy 2131 CTGCTTGTCTTCCAGTATGTTG-CATTTCCAGGCTTATTTGTAATAGGTT----- 2183
Db 74160 TTTACCTTGTGTCAGTTGTTGTTGAAATTTCTCAGGCTGCTGTGAGGTTTGTGG 74101
Qy 2184 TTAGCAGGACATATTGCTCTGTTGT-----TATTGCTGTGTTTGTCTTGGCATA 2237
Db 74100 CTCTACTGGAGACGATGCTCTGCTGTTTGTGTTTATGCTGTGTTTACACTTT 74041
Qy 2238 TAGAGGCTGAGTTGGATGATTTGAATCTAGTCTGAT 2279
Db 74040 TAGGATCTGGTTTGGAAAGTTTGAATTTCTAAGTCTGTT 73999

RESULT 11

AC099619 236604 bp DNA linear HTG 24-JAN-2003
LOCUS Mus musculus clone RP23-465J20, WORKING DRAFT SEQUENCE, 25
DEFINITION unordered pieces.

AC099619

VERSION AC099619.2 GI-27884974

KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Euteleostomi;

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Murinae; Mus.

Birren, B., Nusbaum, C. and Lander, E.

Mus musculus, clone RP23-465J20

Unpublished

2 (bases 1 to 236604)

REFERENCE

AUTHORS

TITLE

JOURNAL

AUTHORS

1 (bases 1 to 236604)

Birren, B., Nusbaum, C. and Lander, E.

Mus musculus, clone RP23-465J20

Unpublished

2 (bases 1 to 236604)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,

Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouckhgalter, B.,

Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,

Choe, Y., Colangelo, M., Collins, S., Collins, S., Collymore, A., Cook, A.,

Cooke, P., DeArillano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,

Ferraira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,

Ginde, S., Gird, S., Goyette, M., Graham, L., Grand-Pierre, N.,

Hagos, B., Headford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,

Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K.,

Lamarez, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,

MacLean, C., MacDonald, P., Majot, J., Marquis, N., Matthews, C.,

McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J.,

Meneus, L., Miheva, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C.,

Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, P.,

Oliiver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,

Ramond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,

Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,

Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,

Straus, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,

Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,

Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,

Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

TITLE

JOURNAL

REFERENCE

1 (bases 1 to 236604)

Submitted (16-NOV-2001) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 236604)

QY 2170 ATTGTAATAGGCTTTAGCAGGACATATTGCTCGGTTGTTATGTTCTGTTGCT 2229
 Db 75425 AGGGATGTTGGCTCTAGTAGGGGCAATTGTTCCAGCTTTATTGACTGTTTAAAT 75484

QY 2230 TTGGCATATAGACGCTCAGTTTGGGATGATTGTAATTTCTAGGTCGTGAT 2279
 Db 75485 CTGATGCTAGGCATCTGGGTTTGGGATGATTGTTAGTTCTAGAGCTGAT 75534

RESULT 12
 AC124411
 LOCUS 177414 bp DNA linear ROD 13-NOV-2003
 DEFINITION Mus musculus BAC clone RP24-267C20 from chromosome 13, complete
 AC124411
 VERSION AC124411.5 GI:29124267
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 177414)
 VanBrunt, A., Bielicki, L., Van Brunt, A. and Kozlowski, A.
 The sequence of Mus musculus BAC clone RP24-267C20
 Unpublished (2001)
 2 (bases 1 to 177414)
 Wilson, R.
 Sequencing of Mus musculus
 Unpublished (2001)
 3 (bases 1 to 177414)
 McPherson, J.D. and Waterston, R.H.
 Direct Submission
 Submitted (14-JUN-2002) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 4 (bases 1 to 177414)
 McPherson, J.D. and Waterston, R.H.
 Direct Submission
 Submitted (06-FEB-2003) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 5 (bases 1 to 177414)
 McPherson, J.D. and Waterston, R.H.
 Direct Submission
 Submitted (19-MAR-2003) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 6 (bases 1 to 177414)
 Wilson, R.
 Direct Submission
 Submitted (13-NOV-2003) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Mar 19, 2003 this sequence version replaced gi:28261652.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu
 Contact: submissions@watson.wustl.edu
 ----- Summary Statistics

 Center project name: M_BB0267C20

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phased quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see http://genome.wustl.edu

SOURCE INFORMATION:

The RP24-24 BAC library has been constructed by Pieter de Jong and coworkers (http://www.chori.org) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at http://www.chori.org

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES
 source
 1..177414
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="13"
 /map="13"
 /clone="RP24-267C20"
 /clone_lib="RP24-24"
 1..48
 /rpt_family="Alu"
 147..249
 /rpt_family="L2"
 391..509
 /rpt_family="L1"
 646..828
 /rpt_family="B2"
 1207..1634
 /rpt_family="L1"
 2321..3214
 /rpt_family="L1"
 2420..2803
 /note="Sequence derived from PCR product of project DNA."
 3391..3798
 /rpt_family="MaLR"
 3799..4204
 /rpt_family="MaLR"
 4819..4916
 /rpt_family="L2"
 4925..5391
 /rpt_family="RMR17C"
 5049..5393
 /rpt_family="ERVK"
 6462..6779
 /rpt_family="MaLR"
 7374..7583
 /rpt_family="B4"
 7877..8013
 /rpt_family="MaLR"
 8215..8336
 /rpt_family="MaLR"
 8696..8914
 /rpt_family="L2"
 9013..9297
 /rpt_family="MaLR"
 9882..10179
 /rpt_family="B4"
 10601..10682
 /rpt_family="L1"
 11927..12038
 /rpt_family="Alu"
 13322..13594
 /rpt_family="B4"
 14823..15002
 /rpt_family="MaLR"
 15218..15612
 /rpt_family="MaLR"
 15613..16707
 /rpt_family="MaLR"
 16708..17102
 /rpt_family="MaLR"

QY 1424 GTAGTCTCTTGGAGTCTGTAGCACATCTGTGAGGGCCCTTCTTACATTTTGGTTCTAT 1483
 Db 152778 GTAGTCTCTTAGATCTGTGAACATCTGTCCAGGCTCTTCTGGCTTTTAGAGTCTCCAG 152837
 QY 1484 TCGAAAGTCAGGTGTAATCTTAATACATCTGCCCTTATATGTTAAATGGTCTTTTTC 1543
 Db 152838 T-GAAGAGTGGGTATAATCTATCAGGTCGCGCTTATATGTACTTTTC-TTTTCCC 152895
 QY 1544 CTTCATCTTTTAATATCTTCTTGTCTATATCTTTAGTGATTCGATTAATATGACAC 1603
 Db 152896 CTTCAGCTTATAATATCTCTTATTTTGTATGTTT-----AGTATTTATTAGTGC 152949
 QY 1604 TGTGGGAGTCTTTTCCGGTCCAACTATTTGGGTGTTTGTATGCTTCTGTACCTTG 1663
 Db 152950 TGAAGGAGCTTTCTTTCTGGTCTAGTCTACTTTTGTGTTCTATATGCTTCTGTACCTTG 153009
 QY 1664 ATAGGCATCTCTTCTCAAGGTTAGGAATTTTCTTTTGGTCTTCTTCTGAAATATTT 1723
 Db 153010 ATAGATGCTCTTT-----TGGTTAGGAAGTTTCTTCTATGATTTTGTGAAATATTT 153065
 QY 1724 TCCCTGCTTTTACCTGCTCTTCTTCCCTTCTCTA-----TTCCTTGGTTTTCGA 1776
 Db 153066 TCTGTGCTGTGTTCTTCTTCTCTCAATTTATATTTACTTTTAGATTTTTTCTG 153125
 QY 1777 TAGTCTCTGCTTCTCTGGATGTTTATGCTGGATTTTATAGCTTACATTTTCTT 1836
 Db 153126 TACTGTCCAAGATCCCTGGAGTTTATGCGAGATTTTATAGACTTAACATTTACT 153185
 QY 1837 TGACCAAGTATFCCATTTCTTCTATCTTGTCTTCACTGCTCCAGATTTCTCTCTATCT 1896
 Db 153186 TGACTAAGTAGTCTTCTTCTTCTATCATGCTTCAATGCCAGAGCTCTCTCTATCT 153245
 QY 1897 CTGTATCTGTGCTAGTGGCTGTCTGAGGTTCTCT-GTTGGGTCTTAAATTTTCA 1955
 Db 153246 CTGTATCTATTTGCTAGGTTTGCCTCTGAGGTTCTCTGTTAAGTTCTCTACATTTTCA 153305
 QY 1956 TTTCAGATTTCTCTCAGTTGGGTTTGTATTAATCTCATTTCCATTTTCAAGTCTCT 2015
 Db 153306 -----AACCCTTAGTCTGGGTTTCTTATGATTTTCTTCCAGTTTGGGTTCT 153356
 QY 2016 GAAATGTTTACTCATTTTCTCCAGTATTTACATTTTCAATTTTCAAGTTTCTTAAATGGA 2075
 Db 153357 CAAGCATATTTTCA-TCTCTGTTGCTGTGTATGTTTTCAGAGATTTCTTTAAGGGGGA 153415
 QY 2076 TATTCATTTCTCTCAAGGACCTTTTATGATTCATTAATGATGTTAGGTCCTTGC 2135
 Db 153416 TATTCATTTCTCTGTAAGGACTTCTATCATATTCATAAAGGTTATTTAAGGCCCATTT 153475
 QY 2136 CTGTGCTTCAGCTATGTTGC-ATTTCTCAGGCGCTTATGTAATAGGGT-----TTTA 2186
 Db 153476 CTGTGCTTCAGCTATGTTACAATACTCAGGAAATGCTGTGTAAGGTTACTGAGCTTA 153535
 QY 2187 GCAGGACATATGTCCTGGTGTGTTATGCTGTGTTTCTTGGCATATAGACGCT 2246
 Db 153536 GTGGAGACATATGTCCTGGCTGTTGTTGATTTGTTATGTTATGTTATGCTAGACCTAGGCA 153595
 QY 2247 GAGTTTGGATGATTTGAATTTCTAGGCTGCTAT 2279
 Db 153596 GCATGTGGAGACTGTAACTTAGTGTCTGAT 153628

RESULT 13

AL672047

LOCUS

DEFINITION Mouse DNA sequence from clone Rp23-316B4 on chromosome X, complete

sequence.

ACCESSION AL672047.9

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 184474)

Leongamornlert, D.

Direct Submission

Submitted (11-SEP-2002) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk

On Sep 17, 2002 this sequence version replaced g1:22265406.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Emi., EMBL; Swi., SWISSPROT; Tr, TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/projects/C_elegans/wormpep RP23-316B4 is from the RPL-23 Mouse PAC library

constructed by the group of Pieter de Jong.

For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACE3.6.

FEATURES

source

1..184474
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosomes="X"
 /clone="RP23-316B4"
 /clone_lib="RPCI-23"

Query Match 32.0%; Score 729.201; DB 1; Length 184474;
 Best Local Similarity 68.7%; Pred. No. 0.097;
 Matches 1295; Conservative 0; Mismatches 508; Indels 82; Gaps 22;

QY 432 AGPATTTTCAATGTCTTTTATCTGTGCGAGACTTGTCTTTTGAATAATGATTTCA 491
 Db 178221 AGATGTTTTCATTTTCTTTCATCTGTAGAGATTTGCTTTTGTGAAGAAATATGTTGTTCA 178280
 QY 492 ATTTTGGAG-AGTTTCATAGGTCGTGACAGAGGACAGTCTTTTGTGTTGGTGA 549
 Db 178281 ATTTTGGAGAAAGTCCATGAGGTCGTAAGTGGAGGCA-TGTTTTATGTTTGTGCGG 178339
 QY 550 ATAGTCTGTAATATCT-CTAGGTCACCTTGGTTTATGACATCAGTACGCTCCAGCATTT 608
 Db 178340 ATGTTCTGTAATATCTGTTAAGTCCATTTGATTTATATATATCAGTTAACTCCAGCATTT 178399
 QY 609 CTCTGTTTCTGTTTGTGAGATGACCTAAGTGTGGAGAAAGTGGGTATTGAAGTAG 668
 Db 178400 TTCTGTTTGTGTTTGGCTGGA-TCAATCCATCTGTGTACTAGTAGGTTATGAAGTTT 178458
 QY 669 CCCACTATCTGTGTGAGGTCATATGTGATTTTAGCTGTAGCTGTGCTGTTTTATGA 728
 Db 178459 CCCACTATCAGTGTGTGAGAGCCAAATATGTGATTTAAGCTAATAGTATTTCTTTTATAA 178518
 QY 729 ACTTGGGTGACATTTGTTTGTGATAGACATTAAGAAATTCATGCTCTCTTGTGGA 788
 Db 178519 ACTTGGGTGCTCTTGTGCAATTTGGTGTAGATGTTTAAAAATGTAAGTCTTCTTGTGGA 178578
 QY 789 TT-TTCTTTTGATGCTATGTAGTATTCTTCCCAATCTCATCTGCTTAGT-----TTTGGG 843

Ds	178579	TTATTCCTTTATGAGTATGAGTTTCTCTCTTATCTCTTATGATGAGTGGGTTTTTT	178638
Qy	844	TTTAAGTCTATTAGTCAGATATAAATGACGTATGCGCTTGTCTCTTAGGGCCATTGG	903
Ds	178639	TGAACCTGCTTTATCTGATATTAATAATCTACACAACTTGCTTAGAGTCCATTGG	178698
Qy	904	CTTGAATATCTTTTCCATCCCTTTTACCTCTAAGTGATGCTATCC--ATGGTAGGTGT	961
Ds	178699	CTTGAACCTACTTTTCCCAACATTTACCTCAGGTGATGCTATCCCTTACGTTAGGATGT	178758
Qy	962	CTTTTTCGGA--TGACAGTAGTAGGATGATCTGTTTTCATATCCATCTGTTACCCAGTA	1020
Ds	178759	ATTCTTGATGTTGCCAGAGAGGGAICTGTTTTTGCAATCCATCTCTTCTGCTGTT	178818
Qy	1021	TCCTTTTCT-----AGAGAAATTAAGATCATTTAGTCATTCATGATTTGAGAAAT	1068
Ds	178819	GTCCTTTT-TAATGGAAATTCAGACCACTGAAGTTGAACAGCAGTCTTGTGTAAT	178878
Qy	1069	ATCAATGAGCAGTGTGTGATTCCTGTTATCTTGACATCTGTGAAGTGTGTGCTGT	1128
Ds	178879	GCTATTTTGTGTTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	178938
Qy	1129	GTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1188
Ds	178939	TGGTAGTAGT	178996
Qy	1189	GCTCTGT	1248
Ds	178997	-----TCCTTTGATTTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	179035
Qy	1249	CATATTTTCTTGAATGCGGTAACAT-CTTTAGATTTGAAGTTTTCCTCTAG--CCTTCT	1305
Ds	179036	TGGGTTTTCATGGGTGATGTTAACTTACTTATTTTGGATTTTTCCTTCTAGTATCTTT	179095
Qy	1306	TTAGTCTGCAATTTGAAGATAGATATCTTTTACATCTGATTTTATCTTAGAATCTTTC	1365
Ds	179096	GTAAGCTGGATTTATAGACAGATATGCTTA-----AATTTTTCATGGAACATCTTAG	179150
Qy	1366	TTTCTCACAATTTGTGACAGAAAGTTTTCCTAAGTCAGTAGCTGCGCTGACATCTGT	1425
Ds	179151	TTTCTACATCTATTATGATGAAAGTCTGTGCGGTATAGTAGCTGAGCTGGCAATTTAT	179210
Qy	1426	AGTCTCTTGGAGTCTGTAGCACATCTGTGACGGCTTCTTACATTTTGAATTTCTATTG	1485
Ds	179211	GGTCTCTTAAAGCTACAGCACATCTGCCAGTCTCTGCGCTTTAAATTTCCATT--	179269
Qy	1486	GAAAGTCAGGTGTAATTTCTAATACATCTGCTTTATATGTTAATGCTCTTTTCCCT	1545
Ds	179270	GAGATCTCAGGCATAATTTCTAATAGTCTGCGCTTTATATGTTAATGCTCTTTTCCCT	179328
Qy	1546	TGCATCTTTTATATCTTTCTTTCTTCTTCTATCTATCTTTTATGATTTATATGCACTG	1605
Ds	179329	TGCACCTTTTATATCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	179384
Qy	1606	TGGGAGTCTTTTCCGGTCCAATCTATTTTGGTGTGTTTGTATGCTCTTCTTGTACCTTGT	1665
Ds	179385	ATGGAAATCTTTCTTCTGCTCTAAT--ATTGATGTTCTGATGCTCTTCTTGTACCTGGAT	179441
Qy	1666	AGGATCTCTTCTCAAGGTAGGAATTTTCTTTTGTGTTTCTTCTTCTTCTTCTTCTTCT	1725
Ds	179442	AGGCATGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	179501
Qy	1726	CCTGTTTGTACCTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1785
Ds	179502	TGTGCTTTTGTACCTGGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	179561
Qy	1786	TGGCTTCTGAGATTTTATGCTGGATATTTTATAGCTTAACATTTTCTTCTTCTTCTTCTT	1845
Ds	179562	AGGCTTCATGATATTTTGTGCCAAGAGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	179621
Qy	1846	TATCCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1905
Ds	179622	TATCCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	179681
Qy	1906	TGTCAGTGGGCTTGTCTCTGAGGTTCCCTGT--TGGGTTCTTAAATTTTCTTCTTCTTCTTCT	1964
Ds	179682	TGCTGTGAGCTTGTCTCTGAGTCTTCTGCTCCCTTTTCAATTTTCTTCTTCTTCTTCTTCT	179741
Qy	1965	TTCCCTCAGTTGGGTTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	2024
Ds	179742	TTTCCCAAGTTTCCATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	179801
Qy	2025	TACTCATTTTCTCCCAAGTATTTACATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	2084
Ds	179802	TATTCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	179859
Qy	2085	CTCTTCTCAAGGACCTTTTATGAATTCATAAATGATGTTAAGTCTTCTTCTTCTTCTTCTTCTT	2144
Ds	179860	-----TTTAAAGGATCCCTATCATATTCATAAAGCTATTTTAAATATTTTCTTCTTCTTCTT	179915
Qy	2145	CAGCTATGTTGCAAT-----CTCAGGCGCTATTTAATAGGTTTCTTCTTCTTCTTCTTCTTCTT	2195
Ds	179916	CATTTATGTTTAACTCCTCAGGACCTGCTGTGGTGGAGTTGTTGGGTTTTTCTTCTTCTTCTTCTT	179975
Qy	2196	TATTCCTCCTGTTTGTATTTCTGCTGTG--TTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	2254
Ds	179976	TACTGCTCCTGCTGTTTATTTGTTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	180035
Qy	2255	GATGATTTGATTTCTAGGTGCTGAT 2279	
Ds	180036	GAAGATTGTAATTTCTAGGTGTTGAT 180060	
RESULT 14			
AC093354			
LOCUS			
DEFINITION			
AC093354 Mus musculus clone RP23-20D11, WORKING DRAFT SEQUENCE, 3 unordered			
pieces.			
AC093354			
VERSION			
AC093354.2 GI:22325146			
KEYWORDS			
HTG: HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.			
SOURCE			
Mus musculus			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;			
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE			
1 (bases 1 to 204031)			
AUTHORS			
Birren, B., Nusbaum, C. and Lander, E.			
TITLE			
Mus musculus, clone RP23-20D11			
JOURNAL			
Unpublished			
REFERENCE			
2 (bases 1 to 204031)			
AUTHORS			
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,			
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B.,			
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,			
Choe, Y., Collangelo, M., Collins, S., Collamore, A., Cook, A.,			
Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,			
Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,			
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,			
Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,			
Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K.,			
Lamazar, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,			
Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,			
McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J.,			
Menues, L., Michova, T., Mieng, V., Murphy, T., Naylor, J., Nguyen, C.,			
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, P.,			
Oliver, J., Peterson, K., Phurkhang, P., Pierre, N., Pollara, V.,			
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,			
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,			
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,			
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,			
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,			
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,			
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.			
Direct Submission			
TITLE			
JOURNAL			
Submitted (20-AUG-2001) Whitehead Institute/MIT Center for Genome			
Research, 320 Charles Street, Cambridge, MA 02141, USA			

REFERENCE 3 (bases 1 to 204031)
 AUTHORS Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
 Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,
 Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
 Cook,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Fero,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
 Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
 Horcon,B., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A.,
 Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
 Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
 McCarthy,M., Melrim,J., Meneus,L., Mihova,T., Mienga,V.,
 Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
 O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
 Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
 Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
 Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
 Testaye,S., Theodore,J., Topham,K., Travers,M., Vassiliiev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
 Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
 Submitted (20-AUG-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Aug 20, 2002 this sequence version replaced gi:15213899.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIGR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
 Center project name: LI3634
 Center Clone name: 20.D.11

----- Summary Statistics
 Sequencing vector: Plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 203365 bases at least Q40
 Consensus quality: 203606 bases at least Q30
 Consensus quality: 203724 bases at least Q20
 Insert size: 194000; agarose-fp
 Insert size: 203831; sum-of-contigs
 Quality coverage: 14.1 in Q20 bases; agarose-fp
 Quality coverage: 13.4 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 121934: contig of 121934 bp in length
 * 121935: Gap of 100 bp
 * 122035: contig of 20006 bp in length
 * 142040: Gap of 100 bp
 * 142041: contig of 61891 bp in length.

FEATURES

source

1. 204031
 /organism="Mus musculus"
 /mol_type="Genomic DNA"
 /db_xref="taxon:10090"
 /clone="RP23-20D11"
 /clone_lib="RPCI-23 Female Mouse BAC"

misc_feature

1. 121934

/note="assembly_fragment"
 clone_end:SP6
 vector_side:left"

misc_feature

122035..142040

/note="assembly_fragment"

misc_feature

142141..204031

/note="assembly_fragment
 clone_end:r7
 vector_side:right"

Query Match 32.0%; Score 725.201; DB 1; Length 204031;
 Best Local Similarity 68.7%; Pred. No. 0.1;
 Matches 1295; Conservative 0; Mismatches 508; Indels 82; Gaps 22;

QY 432 AGTATTTTCAATTCCTTTTATCTGTCGACACTGCTTTGTTGAAATATGATTCACA 491
 DB 177336 AGATGTTTTTCAATTTCTTTTCATCTGTAGAGATTGCTTTTGTGAAGAATATGGGTCA 177395
 QY 492 ATTTTGGAG-AGTTTCATAGGGTGTGACAAAGAGGTACAGTCTTTGTTTGGTGAA 549
 DB 177396 ATTTGGAGAAAGTTCATGAGGTGCTAAATAGGAGCA-TGTTTTATGTTTGTCCGG 177454
 QY 550 ATAGTCTGTAATATCT-CTAGGTCCACTTGGTTTATGACATCAGTTAGCTCCAGCATTT 608
 DB 177455 ATGTTCTGTAATATCTGTTAAGTCCATTTGATTTTATATATATATCAGTTAACTCCAGCATTT 177514
 QY 609 CTCTGTTTCGTTTTTGTGAGATGACCTAATCTGTTGGAGAAATGGGTATTGAAGTAG 668
 DB 177515 TTTCTGTTTTTGTGCTCGA-TCAATCCATCTGTTGACTAGTAGGGTATGAGTTT 177573
 QY 669 CCCACTATCTGTGTGAGGTCAATATGTAATTTAGTGTAGTGTGCTGTTTTATGA 728
 DB 177574 CCCACTATCAGTGTGTGAGAGCAATATGTGATTTAAGCCTAATAGTATTTCTTTATAA 177633
 QY 729 ACTTGGGTGACATTTGTTTGGTGCATAGACATTAAGATTTGCAATGCTCTGTCGGA 788
 DB 177634 ACTTGGGTGCTCTTTCGATTTGTGTCATAGATTTAAAAATTGTAAGTCTTCTTGATGA 177693
 QY 789 TT-TTCCTTTGATGCTATGATGATTTCTTCCCAATCTCATCTGCTAGT-----TTTGGG 843
 DB 177694 TTATTCCTTTTATGAGTAGTAGTTTCTTCTTATCTTATGATTAGTGGGTTTTT 177753
 QY 844 TTTAAGTCTATTAGTCAGATATTTAAATGACATGATCGGCTTCTTAGGGCCATTGG 903
 DB 177754 TGAACCTCTCTTTTATCTGATATTTAAATATCTACACCACTTGCTTCTTAAGTCCATTGG 177813
 QY 904 CTTTGAATATCTTTTCCATCTTCTTACTCTAAGGTGATGTCATCC--ATGGTAGTGTGT 961
 DB 177814 CTTTGAAGTACTTTTCCCAACATTTACCTGAGGTGATGTCATCTTACGTTAGAGT 177873
 QY 962 CTTTTTTTGA-TGCAGCAGTAGATGGAFTTGTGTTTTCATATCCATTCGTTTACCAGTA 1020
 DB 177874 ATTTCTTGATTTGCCAGAGAGGAGGATCTGTTTTTGCATCCATCTCTTTGTCTGT 177933
 QY 1021 TCTTTTTCT------AGAGAAATTAAGATCAATGATGATGTTGAGAATT 1068
 DB 177934 GTCTTTTTTATTGGGAAATTAAGACAGAGTAAGTTGAACAACAGATGCTTGTGTATAAT 177993
 QY 1069 ATCAATGACAGTGTGTTGGGATTTCTGTTATCTTGTGCACTTGTGAAGTGTGTGTGT 1128
 DB 177994 GCTATTTTGTGTGTGTTGTTGCTGCTGCTGTTGTTGTTGTTGTTGTTGTTGTTG 178053
 QY 1129 GT 1188
 DB 178054 TGCTAGTAGT-- 178111
 QY 1189 GTCTGT 1248
 DB 178112 -----TCTTTTGTGTTGCTGGTACTGAGACTATTACTTATCTCT 178150
 QY 1249 CATATTTTCTGAATGTTGGGTAAAT-CTTTAGATTAAGATTTTCTCTCTAG--CCTTCT 1305
 DB 178151 TGGGTTTTTCATGGTGTAGTTAACTTACTTTATTTTGGATTTTTTCTTCTAGTATCTTTT 178210
 QY 1306 TTAGGTCTCATTTTGAAGATAGATATTCTTTACATCTGATTTTATCTTAGAATGCTTTTC 1365
 DB 178211 GTTAGGCTGGATTTATAGACATATATGCTTA-----AATTTTTCAAGAACATCTTAG 178265
 QY 1366 TTTCTCCAACTATTGTGACAGAAAGTTTTTCTAAGTGAAGTAGTCTGGCCTGACATCTGT 1425

QY 1864 TGCTTCTACTGCTGAGATTCCTCTTCTATCTCTTGTATCTCTGTCAGTGAGGCTTGCTGTCT 1923
 Db 18076 TGTCTTTAACAGCTGAGATTCCTTCTTCCATATCTGTCATCTGTCGTGTCGTGTCGTAT 18135
 QY 1924 CTGAGGTCCTGCTGGTT-CTTAATTTTTCATTCAGATTTCCCTTCAGTTGGGTTT 1982
 Db 18136 CTGTAGCTCCGTCCTACCTACTAGATTTTCACTTCATGTCCTTCATTTATGTTT 18195
 QY 1983 TGTTTTAAATCTATTTCCACTTCAGGTCCTGAATGTTTACTCATTTCTCCACG 2042
 Db 18196 TCTTTATGCTTTTATTTCCATTTTCAGTCTTGAACAGTTTACTCATTTCTTCACT 18255
 QY 2043 TATTTAC--ATTTCATAGGTTCTTTAATGATTTATTCATTTCCCTTCAGGACCTT 2100
 Db 18256 GTTGTCTGCAATTTCTGGATTTTCAAGGATTTATTTGTTTCCCTTTAAGGCTTC 18315
 QY 2101 TTATGAATTCATAAATGATGTTAAGTCTTGGCTTGTCTTCAGCTATGTTGCATTC 2160
 Db 18316 TATCACCTTCATAAGTGTGTTTAAAGTTATTTACTTGTTCAGCTGTGTTGGAATG 18375
 QY 2161 TCAGGGCC-----TATTGTAATAGGTTTATAGCAGGACATATTCCTCGTTGTT 2211
 Db 18376 TCCAGAGCTGCTGCTGATGATAGTTAAGTTCGGGTGATACCATGTTGCCAAGCTATT 18435
 QY 2212 ATTGTCCTGTTTTCCTTGGCATATAGACGGTGGAGTTGGGATGATTTGTAATCTAG 2271
 Db 18436 GTTGTATGTTGTTTATCTGCTGCTTCAAGTACCTGATTTGGATTAATATAGGTTTAG 18495
 QY 2272 GTGCTGAT 2279
 Db 18496 GTACCAAT 18503

RESULT 17
 AC115852/c
 LOCUS AC115852 142529 bp DNA linear HTG 21-OCT-2002
 DEFINITION Mus musculus clone RP24-270L8, WORKING DRAFT SEQUENCE, 6 ordered
 pieces.
 ACCESSION AC115852
 VERSION AC115852.4 GI:24182195
 KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 Birren,B., Nusbaum,C. and Lander,E.
 Mus musculus, clone RP24-270L8
 Unpublished
 2 (bases 1 to 142529)
 1 (bases 1 to 142529)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 1 (bases 1 to 142529)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
 Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
 Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
 Cook,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Horton,L., Hulme,W., Illiev,I., Johnson,R., Jones,C.,
 Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
 Lander,A., Lehoczyk,J., Levine,R., Lindblad-Toh,K., Liu,G.,
 Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
 Mihova,T., Miengwa,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strausman,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,D.,
 Topham,K., Travers,N., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zairoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission

JOURNAL REFERENCE AUTHORS

Submitted (22-MAR-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 142529)
 Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
 Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
 Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
 Cook,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
 Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
 Hagos,B., Horton,L., Hulme,W., Illiev,I., Johnson,R., Jones,C.,
 Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
 Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J.,
 Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
 Miengwa,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
 Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
 Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
 Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
 Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
 Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,D., Topham,K.,
 Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
 Wyman,D., Young,G., Zairoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (21-OCT-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Oct 21, 2002 this sequence version replaced Gi:21622796.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

TITLE JOURNAL COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L24684
 Center clone name: 270_L8
 ----- Summary Statistics
 Sequencing vector: Plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 141053 bases at least Q40
 Consensus quality: 141891 bases at least Q30
 Consensus quality: 141937 bases at least Q20
 Insert size: 145000; agarose-fp
 Insert size: 142029; sum-of-contigs
 Quality coverage: 9.5 in Q20 bases; agarose-fp
 Quality coverage: 9.7 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 6 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 1173: contig of 1173 bp in length
 * 1174 1273: gap of 100 bp
 * 1274 4401: contig of 3128 bp in length
 * 4402 4501: gap of 100 bp
 * 4502 99416: contig of 94915 bp in length
 * 99417 99517: gap of 100 bp
 * 99517 113174: contig of 13658 bp in length
 * 113175 113274: gap of 100 bp
 * 113275 140154: contig of 26880 bp in length
 * 140155 140254: gap of 100 bp
 * 140255 142529: contig of 2275 bp in length.
 Location/Qualifiers
 1. 142529
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"

FEATURES source


```

Db      129264  TATCACCTTCATAGGTGGTTTAAAGTTATTTACTTGTTGTTTCACCTGTGTGGAATG 129205
Qy      2161   TCAGGGCC-----TATTGTAATAGGGTTTACAGGACATATTCCTGGTTGTT 2211
Db      129204  TCCAGAGCTTCTGTAGTAGGATAGTTAAGTTCGGGTGATACCATGTTGCCAAGGCTATT 129145
Qy      2212  ATTGCTCTGTTTTCCTTTGGCATATAGACGGCTGAGTTTGGGATGATGTAATCTAG 2271
Db      129144  GTTGATTGGTTCCTATGCTGGCTCTAAGTACCTGGATTGGGATATATAGGTTAG 129085
Qy      2272  GTGTGTAT 2279
Db      129084  GTACCAAT 129077

RESULT 18
AC129304      153772 bp      DNA      linear      ROD 27-NOV-2003
DEFINITION   Mus musculus BAC clone RP24-462H23 from chromosome 10, complete
              sequence.
ACCESSION   AC129304
VERSION     AC129304.6 GI:30425602
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
REFERENCE   1 (bases 1 to 153772)
AUTHORS     Shah,N. and Meyer,R.
TITLE       The sequence of Mus musculus BAC clone RP24-462H23
JOURNAL     Unpublished (2001)
REFERENCE   2 (bases 1 to 153772)
AUTHORS     Wilson,R.
TITLE       Sequencing of Mus musculus
JOURNAL     Unpublished (2001)
REFERENCE   3 (bases 1 to 153772)
AUTHORS     McPherson,J.D. and Waterston,R.H.
TITLE       Direct Submission
JOURNAL     Submitted (28-JUL-2002) Genome Sequencing Center, 4444 Forest Park
              Parkway, St. Louis, MO 63108, USA
REFERENCE   4 (bases 1 to 153772)
AUTHORS     McPherson,J.D. and Waterston,R.H.
TITLE       Direct Submission
JOURNAL     Submitted (15-MAR-2003) Genome Sequencing Center, 4444 Forest Park
              Parkway, St. Louis, MO 63108, USA
REFERENCE   5 (bases 1 to 153772)
AUTHORS     McPherson,J.D. and Waterston,R.H.
TITLE       Direct Submission
JOURNAL     Submitted (08-MAY-2003) Genome Sequencing Center, 4444 Forest Park
              Parkway, St. Louis, MO 63108, USA
REFERENCE   6 (bases 1 to 153772)
AUTHORS     Wilson,R.
TITLE       Direct Submission
JOURNAL     Submitted (27-NOV-2003) Department of Genetics, Washington
              University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
              On May 8, 2003 this sequence version replaced gi:28973954.
COMMENT     ----- Genome Center
              Center: Washington University Genome Sequencing Center
              Center code: WUGSC
              Web site: http://genome.wustl.edu
              Contact: submissions@watson.wustl.edu
              ----- Summary Statistics
              Center project name: M_BB0462H23
              -----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate

chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclones; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RP24-24 BAC library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AC122018.

FEATURES	source
repeat_region	1. 153772 /organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090" /chromosome="10" /clone="RP24-462H23" /clone_lib="RP24-462H23"
repeat_region	326. 459 /rpt_family="Alu"
repeat_region	546. 747 /rpt_family="MIR"
repeat_region	1790. 1966 /rpt_family="MaLR"
repeat_region	1971. 2070 /rpt_family="B4"
repeat_region	2644. 2685 /rpt_family="B4"
repeat_region	2782. 2964 /rpt_family="MER1_type"
repeat_region	3483. 3586 /rpt_family="MIR"
repeat_region	3648. 3768 /rpt_family="Li"
repeat_region	3751. 4429 /rpt_family="Li"
repeat_region	5506. 6144 /rpt_family="Li"
repeat_region	6425. 6643 /rpt_family="ERVK"
repeat_region	6772. 6818 /rpt_family="ERVK"
repeat_region	6862. 6953 /rpt_family="Alu"
repeat_region	6958. 7146 /rpt_family="ERV1"
repeat_region	7886. 8253 /rpt_family="MaLR"
repeat_region	8533. 8916 /rpt_family="MaLR"
repeat_region	9222. 9364 /rpt_family="Li"
repeat_region	9365. 9665 /rpt_family="MaLR"
repeat_region	9666. 9692 /rpt_family="Li"
repeat_region	10041. 10071 /rpt_family="Li"
repeat_region	10086. 10673 /rpt_family="Li"
repeat_region	11040. 11175 /rpt_family="Li"


```
repeat_region /rpt_family="MIR"
11755..11894
/rpt_family="B4"
11944..12052
/rpt_family="MaLR"
12105..12345
/rpt_family="MaLR"
12670..12786
/rpt_family="L1"
12801..12913
/rpt_family="MaLR"
13069..13137
/rpt_family="ID"
13451..13916
/rpt_family="ERV"
14644..14697
/rpt_family="ERV1"
15524..15662
/rpt_family="L1"
15715..16257
/rpt_family="L1"
16258..16391
/rpt_family="Alu"
16271..16414
/rpt_family="B4"
16427..16767
/rpt_family="L1"
18746..19268
/rpt_family="RMR1A"
19290..19366
/rpt_family="L1"
19427..21862
/rpt_family="L1"
21865..22790
/rpt_family="ERVX"
23580..25720
/notes="Unresolved tandem repeat."
repeat_region 25844..25980
/rpt_family="L1"
26029..26087
/rpt_family="ERV1"
27269..27529
/rpt_family="MaLR"
27541..27672
/rpt_family="Alu"
29144..29254
/rpt_family="L2"
30430..30823
/rpt_family="MaLR"
30874..30966
/rpt_family="B4"
31087..31369
/rpt_family="L1"
31824..31916
/rpt_family="L1"
32281..32572
/rpt_family="MaLR"
32825..32874
/rpt_family="L1"
32885..32997
/rpt_family="L1"
32984..33289
/rpt_family="L1"
33863..33953
/rpt_family="ERV1"
33954..34109
/rpt_family="MaLR"
34259..34465
/rpt_family="MaLR"
34666..34536
/rpt_family="ERV1"
34543..34743
```

```
Query Match 31.4%; Score 715.001; DB 1; Length 153772;
Best Local Similarity 68.1%; Pred. No. 0.1;
Matches 1272; Conservative 0; Mismatches 490; Indels 106; Gaps 21;

QY 270 TTCCCTCTCTCTCCAAACACACTCTATTCTTTGATTTCTATCTTGCTC-ATTTTAACT 328
Db 20081 TTATTCAATCCTAGGAAGTCCTTTCTTTTAAATTTCTGCCCTGACACTTTTTCATT 20140
QY 329 CAGTAGTGAGTTGTTTGGTTTCCATAAGTTTGTAGTTTTCGTGTTCTCTGTGTGTT 388
Db 20141 GAGTAGAGAGTTGTTTCAGT-----GTCTACACTTTCTGTGTTCTATT 20184
QY 389 GTTGTATTCTAGATTTAAGCTGTGGTGGTCAGATAGACATAGATATTATTCAATTCT 448
Db 20185 ATTGATATCCAGCTTTAATCCATGGTGGTCTGATAGGGTGCGAGGAGTTATTTCACATTT 20244
QY 449 CTTTATCTGTGAGACTTGTCTTTGTTTGAATAATGTAATCAATTTTCGAG--AGTTTC 506
Db 20245 CTTGTATCTGCTGAGACTTGTCTTTGTCTGAATATGATCAGTTTTCGAGACAATTC 20304
QY 507 ATAGGGTGTGACAGAAGGTACAGTC-TTTCGTGTTTGGTGAATAGTCTGTAATATC 565
Db 20305 ATGAGGTGCTGAGAGAAGGTATATTCTTTGTGTTGGTGGAAATGTTTGTAGATATC 20364
QY 566 T-CTAGGTCCACTTGGTTTATGACATCAGTTAGCTCCAGCATTTCTCTGTTTCTGTTTTT 624
Db 20365 TGTGGGTTTCATCTGTTTATAACATCAGTTAGCTCCAGCATTTCTCTACTTAGTTTTT 20424
QY 625 GTT-----GAGATGACCTAACTGTTGGAGAGAAATGGGTATTTGAAGTACCCACTATCT 578
Db 20425 TTTTTTTCTGTTGCCCTGCTCTATTGGCAAGTAGGGTATTGAAGTCTCCCACTGTCA 20484
QY 679 GTGTGTG-AGGTCAATATGTGATTTTAGCTGTAGCTGTCTTTGTTTATGAACCTGGGTG 737
Db 20485 CTTGTGAAAGTCAATGCAATGATTTAAGCTACATAGTATCTGCTTATGAACCTGGGTA 20544
QY 738 ACATTGTGTTGGTGCATAGACATTAAGAAATGCAATGTCCTCTTGTTGGAT-TTTCCTT 796
Db 20545 CTCTTGTGTTTGGGCGATAGGTGTTAAGAAATCAAAACGCCATGTTGGTGGATCTTTCCCT 20604
QY 797 TGATGCTATGTAGTATTCTTCCCAATCTCACTGCTTAGTTTGGGTTTAACTCTA-TT 855
Db 20605 TGATGCTGTAAATACCCCTGCCCTGCTCTCTGTTAGTTTGGGTTTGTGTTGATTT 20664
QY 856 AGTCAGATATTAAATGACTGTATCGGCTTGCTTCTTAGGGCCATTTGTTAGATATCT 915
Db 20665 TGTACAGACAGAAATGGCAACACCAGATTGCTTGTAGACTATCTCGGTAGATACCT 20724
QY 916 T-TTCCATCCTTTTACTCTAAGTGATGCTATCCATGCTAGTGTCTTTTGTGATGC 974
Db 20725 TCTTGCAACCTTTTATCTTAGGTAAATGGCTACCTTGAATGAGTAAGTTTGTGA 20784
QY 975 AGCAGTAGGATGGATCTTGTGTTTTCATATCCATCTGTTACCCAGTATCTTTTCTAGAGA 1034
Db 20785 TGCAGCAGAAATGAATCCCTGTTTACATCCCTTCTGTAGACTGTATCTTTTATGGGT 20844
QY 1035 AATTAGATCATTTAGTCAATGATGAGAAATATCAATAGACAGTGTGTTGTGATCT 1094
Db 20845 AATTGAGA-----CCATTGATGTCAGAGATATCAATGAGAAATGGTTTGTGTTCT 20896
QY 1095 TGTATCTTGCACTGTGGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1154
Db 20897 TGTGATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 20956
QY 1155 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1214
Db 20957 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 20999
QY 1215 TCTTTTGATTTTGGCCCGAATTAATTATTATTTATTTTCTTGAATGTGGTAACT 1274
Db 21000 TCTGGGATTAATTAATTCACAGTTTGTGTTTGTGTTTTTTTTCACCTTTGGTGGGGTAACT 21059
QY 1275 CTTTAGATTGAAGTTTCTCTCTA---GCCCTCTTTAGGTCTGCAT-TTGAAGATAGATA 1330
```

```

Db      21060 CTTTAGACTGCAATTAACCTTCCAAAGTGAATCTGTAGGGCTCATCTATTATATAATAA 21119
Qy      1331 TTCCTTACATCTGATTTATCTTAGAATGCTTTCTTCCCACTATTGTGACAGAAG 1390
Db      21120 TTGCTTACATTTGGCTTTATCATAAATGCTGTTTCTTCATCTATGTGATGAAG 21179
Qy      1391 TTTTCTAAGTCAGTAGCTGGCTGACATCTAGTCTCTCTGGAGTGTGTAGCACATC 1450
Db      21180 TTTTCTAGGTATATAGTAGGCTGGCACTGTGTCTCTTAAATCTGTAGCACAT 21239
Qy      1451 TGTGAGGCGCTTCTTACATTTTGAATTTCTATTTGGAAGTCAAGTGTAACTTATATAC 1510
Db      21240 TGTCTAGGCTCTTTTGTGTTAGAGTCTC-----CATGAGAAGTCTGTATAG 21287
Qy      1511 ATCTGCCCTTATATGTTAAATGGTCTTTTTCCTTGCATCTTTTAAATCTTTCTTTG 1570
Db      21288 GTCTGCCCTTATATGTTACTTGG-CCTTTTCCTTAAAGCTTTTAACTTTCTTTG 21346
Qy      1571 TTCATATCTTTTAGTAGATTGATTAATGACACTGTGGGGAGTTCCTTTCCGCTCCAAT 1630
Db      21347 -TAITGTCAGTCAGTGTCTTGTGATAATATATGCTAAAGGGGCTTTCTTCTGCTTCAAT 21405
Qy      1631 CTATTTGGTGTGTTGTATGCTTCTTGATCTTGTACCTGTAGGCATCTCTTCTCAAGGTTAGGA 1690
Db      21406 CTCCTTGTATCTAAATGC-----TACCCTTCTAGGCATCTCTTCTTTTAGTTAGGA 21459
Qy      1691 AATTTTCTCTTTTGGTTTCTTGAAATAATTTTCCCTGCTTTTGAACCTGCTTTCTTCCC 1750
Db      21460 AATTTTCTCTATGATTTTGTGAAATAATTTCTATGACTTTTGAACCTGCTTTTCCCC 21519
Qy      1751 C-----TTCCTCTATCTCTTGGTTGTTGATAGTCTCTGCTTCTGCTGATG 1799
Db      21520 CCCCCCTCTCTATCTCTATTCCTTAGATTAGTCTTTCGATCCCATATTTCTGCTGATG 21579
Qy      1800 TTTTATGCTGCTGATTTTATAGACTTAACATTTTCTTTTGAACCAAGGATTCATTTCTTCT 1859
Db      21580 TTTTATCATCAGGGTCTCTTAGAGTTAAACATTTCTCTAACCAAGGATTCATTTTCT 21639
Qy      1860 ATCTTGTCTTACCTGCTGAGATCTCTCTCTCTATCTCTGTTATCTGTCTGCTGAGGCTT 1919
Db      21640 ACTGTGCTTCTCAGTCTTAAACATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 21693
Qy      1920 GTCTCTGAGGTCCTGTTGGTCTTAAATTTTCTTCTTCAATTTCCAGATTTCTCTTCTGTTGG 1979
Db      21694 -----GTTTCTTACAGTGTCTTAAATTTCTCTTCTGTTTCAAGATTTCTCTTCTGTTGG 21745
Qy      1980 TTTTCTTATTAATTTCTATTTCCACTTTCAGGCTCTGAAATGTTTTACTCATTTTCTCTCC 2039
Db      21746 TTTTCTTATTTAGTTGTTATTCACATTCAGATCTGAACTGTTTCTTCTCA-TTTCTCTCC 21804
Qy      2040 CAGTATTTACATTTTCTAGTCTTCTTAAAGGATTTATTCATTTCTCTTCTCAGGACCT 2099
Db      21805 CACTGTTTTTGTCTTATAGATTTCTTTTGGTAGATTTACTCACCTTCTCTTTTAAAGACCA 21864
Qy      2100 TTTATGAA 2107
Db      21865 GTCAGGCA 21872

RESULT 19
AC122210
LOCUS
DEFINITION Mus musculus chromosome UNK clone RP23-96G7, WORKING DRAFT
SEQUENCE, 23 unordered pieces.
ACCESSION AC122210
VERSION AC122210.2 GI:22475987
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 271609)

```

```

McPherson,J.D. and Waterston,R.H.
The sequence of Mus musculus clone
Unpublished
2 (bases 1 to 271609)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (23-MAY-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 271609)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (25-AUG-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
On Aug 25, 2002 this sequence version replaced gi:21105042.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@watson.wustl.edu
----- Project Information -----
Center project name: M_BA0096G07
----- Summary Statistics -----
Sequencing vector: M13; 0%
Chemistry: Dye-primer ET; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 264453 bases at least Q40
Consensus quality: 266334 bases at least Q30
Consensus quality: 267385 bases at least Q20
Insert size: 234000; agarose-fp
Insert size: 268963; sum-of-contigs
Quality coverage: 15.12 in Q20 bases; agarose-fp
Quality coverage: 11.17 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 23 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1097: contig of 1097 bp in length
* 1098 1397: gap of unknown length
* 1198 1253: contig of 56 bp in length
* 1254 1353: gap of unknown length
* 1354 1743: contig of 390 bp in length
* 1744 1843: gap of unknown length
* 1844 2940: contig of 1097 bp in length
* 2941 3041: gap of unknown length
* 3041 4209: contig of 1169 bp in length
* 4210 4309: gap of unknown length
* 4310 5593: contig of 1284 bp in length
* 5594 6894: gap of unknown length
* 6895 8293: contig of 1201 bp in length
* 8293 8392: gap of unknown length
* 8393 9864: contig of 1298 bp in length
* 9865 9965: gap of unknown length
* 9965 11951: contig of 1472 bp in length
* 11952 12051: gap of unknown length
* 12052 13101: contig of 1050 bp in length
* 13102 13201: gap of unknown length
* 13202 14843: contig of 1642 bp in length
* 14844 17288: gap of unknown length
* 17129 17228: contig of 2185 bp in length
* 17229 18762: gap of unknown length
* 18763 18862: contig of 1534 bp in length
* 18862: gap of unknown length

```

```

* 18863 25092: contig of 5230 bp in length
* 25093 25192: gap of unknown length
* 30277: contig of 5085 bp in length
* 30377: gap of unknown length
* 30378 47187: contig of 16810 bp in length
* 47188 47287: gap of unknown length
* 47288 63647: contig of 16360 bp in length
* 63648 63748: gap of unknown length
* 63748 83396: contig of 21649 bp in length
* 83397 85496: gap of unknown length
* 85497 119995: contig of 34499 bp in length
* 119995 120096: gap of unknown length
* 120096 147799: contig of 27704 bp in length
* 147800 147899: gap of unknown length
* 147900 194953: contig of 47054 bp in length
* 194954 195053: gap of unknown length
* 195054 271609: contig of 76556 bp in length.
FEATURES
    source
        1. 271609
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            /chromosome="UNK"
            /clone="RP23-96G7"
            1. 1097
                /note="assembly_name:Contig14"
            1198..1253
                /note="assembly_name:Contig15"
            1354..1743
                /note="assembly_name:Contig16"
            1844..2940
                /note="assembly_name:Contig28"
            3041..4209
                /note="assembly_name:Contig42"
            4310..5593
                /note="assembly_name:Contig43"
            5694..6894
                /note="assembly_name:Contig45"
            6995..8292
                /note="assembly_name:Contig49"
            8393..9864
                /note="assembly_name:Contig52"
            9965..11951
                /note="assembly_name:Contig53"
            12052..13101
                /note="assembly_name:Contig56"
            13202..14843
                /note="assembly_name:Contig59"
            14944..17128
                /note="assembly_name:Contig61"
            17229..18762
                /note="assembly_name:Contig62"
            18863..25092
                /note="assembly_name:Contig63"
            25193..30277
                /note="assembly_name:Contig64"
            30378..47187
                /note="assembly_name:Contig65"
            47288..63647
                /note="assembly_name:Contig66"
            63748..85396
                /note="assembly_name:Contig67"
            85497..119995
                /note="assembly_name:Contig68"
            120096..147799
                /note="assembly_name:Contig69"
            147900..194953
                /note="assembly_name:Contig70"
            195054..271609
                /note="assembly_name:Contig71"

```

Query Match 31.4%; Score 715.001; DB 1; Length 271609;
 Best Local Similarity 68.1%; Pred. No. 0.16;

```

Matches 1272; Conservative 0; Mismatches 490; Indels 106; Gaps 21;
QY 270 TTCCCTCTCTCTTCAAAACACTTCTATTTCTTGATTTCTATCTTGCTC-ATTTTTAACT 328
DB 143362 TTTATTCAATCCTAGGAAGTCTTTCTTTTTTAAATTTCTGCTTGACACTTTTTTTCATT 143421
QY 329 CAGTAGTAGTTGTTTGGTTTCCATAAGTTTGTAAAGTTTCTGTTGTTCTCTGTTGTT 388
DB 143422 GAGTAGAGAGTTGTTTCAGT-----GTCTACACTTCTCTGTTTCTTCTATT 143465
QY 389 GTTGTATCTAGATTTAAGCTGTGGTGCAGATAGGACATAGAGTATTATTTCATTTGT 448
DB 143466 ATTGATATCCAGCTTTAATCCATGGTCTCATAGGGTGCAGGGAGTTATTTCACATT 143525
QY 449 CTTTATCTGTGCGAGACTTGTCTTTTGAATAATGATTCATTTTGGAG--AGTTTC 506
DB 143526 CTTGTATCTGTGAGACTTGTCTTTGTCTGAATATGATCATGATTTTGGAGACAATTC 143585
QY 507 ATAGGTGCTGACAAGAAGGTACAGTC-TTTCGTGTTTGGTGAATAGTCTCTAAATATC 565
DB 143586 ATGAGGTGCTGAGAAGAGGTATATTTCTTTGTGTTGGTGGATGTTTGTAGATATC 143645
QY 566 T-CTAGTCCACTTGGTTTATGACATCAGTTAGTCCAGCATTTCTGTGTTTCGTTTTTT 624
DB 143646 TGTGGGTTCACTCTGTTTATAACATCAGTTAGTCCAGCATTTCTCTACTTAGTTTTTT 143705
QY 625 GTT-----GAGATGACCTAACTGTTGGAGAAATGGGGTATTGAAGTAGCCCATATCT 678
DB 143706 TTTTTTTCTGGTGGCCCTGCTATTGGCAAGTAGGGTATTGAAGTCCCATCTGCA 143765
QY 679 GTGTGTG-AGGTCAATATGTGATTTTAGCTGTAGCTGTGCTGTTTATGAACCTGGGTG 737
DB 143766 CTGTGTGAAGTCAATGCATGATTTTAAGCTAACATAGTATCTGCTTTATGAACCTTGGGTA 143825
QY 738 ACATTGTGTTTGGTGCATAGACATTAAGAATTGCAATGTCCTCTTGGTGGAT-TTTCCTT 796
DB 143826 CTCTTGTGTTTGGGCATAGGTGTTAAGAATCAAAACGCCATGTTGGTGGATCTTTCCTT 143885
QY 797 TGATGGCTATGTAGTATTCTTCCCAATCTCATCTGCTAGTTTGGTTTAAAGTCTA-TT 855
DB 143886 TGATGGCTGAAATACCTGCGCTGCTCTCTCTGTTAGTTTGGTGTGATGGAATTT 143945
QY 856 AGTCAGATTTAAATGACTGTATCGGTTGCTTCTTAGGGCCATTTGCTTGAATATATCT 915
DB 143946 TGTCAACAGCAAAATGGCAACACAGATTGCTTGTAGGACTATCTCGTGAATACTC 144005
QY 916 T-TTCCATCCTTTTACTCTAAGGTGATGTCATCCATCTGTTAGCTGTTCTTTTTCGATGC 974
DB 144006 TCTTGCAACCTTTTATCCTTAGGTAATGGCTAACCTTGAATGGAATAGGTTTGTGA 144065
QY 975 AGCAGTAGATGGATCTTGTGTTTTCATATCCATTCTGTTACCCAGTATCTTTTCTAGAGA 1034
DB 144066 TGCAGCAGAATGAATCCTGTTTATCATCCTTTCGTTAGACTGTATCTTTTATTGGGT 144125
QY 1035 AATTAAGATCATTTAGTCAATGATTTGAGAATTAATCAATGAGCAGTGTGTTGGATTTCT 1094
DB 144126 AATTGAGA-----CCATTGATGTTGAGATATCAATGAGAATGGTTTGTATCT 144177
QY 1095 TGTATCTTGCACTGTGGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1154
DB 144178 TGTGATTTTGGTGGTGTGATGCGGGTGGGGTGTGTGTGTGTGTGTGTGTGTGTGT 144237
QY 1155 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1214
DB 144238 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 144280
QY 1215 TCTTTTGAATTTTGGCTGGAATTTATTATTATTTTCTTGAATGTGGGTAAACAT 1274
DB 144281 TCTGGGATTTATTATTTCCAGTTTGTGTTTGTGTTTGTGTTGTTGTTGTTGTTGTT 144340
QY 1275 CTTTAGATTGAAGTTTCTTCTCTA---GCCTCTTTTAGTCTGCAAT-TTGAAGATAGATA 1330
DB 144341 CTTTAGACTGGAATTAACCTTCCAAGTACTTCTGTAGGGCTCTATCTATTATATAAATA 144400

```

QY 1331 TTCCTTACATGATTTTATCTTGAATGCTCTTCTTCTCCAACTATTGTGACAGAAG 1390
Db 144401 TTGCTTACATTTGGCTTTATCATAAATGCTCTTTCTTCATCTATTGTGATTGAAG 144460
QY 1391 TTTTCTAAGTCAGTAGCTGGCTGACATCTGTAGTCTCTTGGAGTCTGTAGCACATC 1450
Db 144461 TTTTGTAGGTATATAGTAGTGTAGCTGGCATCTGTCTCTTTAAATCTGTAGCACAT 144520
QY 1451 TGTGAGGCGCTCTTACATTTTCTAGTTTCTTATTTGGAAGTCAAGTGTATTTCTTAATC 1510
Db 144521 TGTCTAGGCTCTTTTGTGTTAGATCTC-----CATGAGAGTCTGTATG 144568
QY 1511 ATCTGCTTTATATGTTAAATGCTCTTTTTCCTTGGCACTTTTAAATCTTTCTTTG 1570
Db 144569 GTCGCTCTTATATGTTACTTGG-CTTTTTCCTTAAAGGCTTTTAAATCTTTCTTTG 144627
QY 1571 TTTCTATCTTTTGTAGTATTCATTTATGATTTATGCTGAGGAGTCTTTCTTCCGCTCAAT 1630
Db 144628 -TAATGTCAGTCAGTGTGTTTGTATTAATTAATGCTGAAGGGGCTTTCTTTCTGTTCAAT 144686
QY 1631 CTATTTGTGTTTGTATGCTTTCTGTACCTTGTAGTCATCTCTTTCTCAAGTGTAGGA 1690
Db 144687 CTCCTTTGTTTCTAAATGC-----TACCTTTCTAGGCATCTCTTCTTTAGTGTAGGA 144740
QY 1691 AATTTTCTTTTGGTTTCTTGAATATTTTCCCTGCTTTTGTAGCTGCTTCTTCTCC 1750
Db 144741 AATTTTCTTCTATGATTTTGTGAATATTTTCTATGACTTTTGTAGCTGCTTCTCC 144800
QY 1751 C-----TTCCTCTATTTCTTGGTTTGTGATAGTCTCTGCTTCTCTGATG 1799
Db 144801 CCCCCCTCTATTTCCATTTATCTTAGATTTAGTCTTTGCATCCCATATTTCTGATG 144860
QY 1800 TTTTATGCTGATTTATTTTATGATTTTATGATTTTCTTTTGTAGCAAGGATCTCTTCT 1859
Db 144861 TTTTATCATCAGGGTCTCTTAGATTTTACACTTTTCTTAACCAAGGTACCAATTTTCT 144920
QY 1860 ATCTGCTCTTCACTGCTGAGATCTCTCTCTATCTCTGTTATCTGTCAGTGAAGCTT 1919
Db 144921 ACTGCTGCTTCACTGCTGAGATCTCTCTCTATCTCTGTTATCTGTCAGTGAAGCTT 144974
QY 1920 GTCTCTGAGTCTCTGTTGGTCTTAAATTTTTCATTTTCAGATTTCTCTCAAGTGTGG 1979
Db 144975 -----GTTCTTTTACCAGTGTGTTAAATTTCTGTTTCAAGATTTCTCTCAAGTGTGG 145026
QY 1980 TTTTGTATTAATCTATTTTCCATTTCCAGTCTGAGTCTGAGTGTATTTCTCAATTTCTCTCC 2039
Db 145027 TTTTGTATTAATCTATTTTCCATTTCCAGTCTGAGTGTATTTCTCAATTTCTCTCC 145085
QY 2040 CAGTATTTTACATTTTTCATGATTTTCTTTTAAATGATTTTATTTCTCTTCTCAAGGACCT 2099
Db 145086 CACTGTTTGTCTTATGATTTCTTTTGGTATTTTACTCACATTTCTCTTTTAAAGACCA 145145
QY 2100 TTTATGAA 2107
Db 145146 GTCAGGCA 145153

RESULT 20
AC126868/c
LOCUS Rattus norvegicus clone CH230-50H7, WORKING DRAFT SEQUENCE.
DEFINITION Rattus norvegicus clone CH230-50H7, WORKING DRAFT SEQUENCE.
ACCESSION AC126868
VERSION AC126868.4 GI:30521522
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 247119)
Muzny,D.Marie., Mettaker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,

Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escoto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gregg, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guera, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Liu, J.,
Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshewa, L., Louisedge, H., Lozano, R.J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E.,
Mawney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwaelele, O., Okwionu, G., Olariunpasegon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plummer, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.,
Puzo, M., Quinlan, J., Rachlin, E., Reeves, K., Richards, S., Riggs, F.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Snajd, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valdes, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, J., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.
Direct Submission
Unpublished
2 (bases 1 to 247119)
Worley, K.C.
Direct Submission
Submitted (10-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 247119)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 10, 2003 this sequence version replaced gi:24819181.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature

table.

```

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GPDW
Center clone name: CH230-50H7
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 235042 bases at least Q40
Consensus quality: 237296 bases at least Q30
Consensus quality: 238799 bases at least Q20
Estimated insert size: 245174; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
  (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
  * consists of 1 contigs. Gaps between the contigs
  * are represented as runs of N. The order of the pieces
  * is believed to be correct as given, however the sizes
  * of the gaps between them are based on estimates that have
  * provided by the submitter.
  * This sequence will be replaced
  * by the finished sequence as soon as it is available and
  * the accession number will be preserved.
  *
  * 1 247119: contig of 247119 bp in length.
  Location/Qualifiers
    1..247119
      /organism="Rattus norvegicus"
      /mol_type="genomic DNA"
      /db_xref="taxon:10116"
      /clone="CH230-50H7"
  1..1116
    /note="wgs end_extension
clone_end:T7"
  2169..2761
    /note="clone_boundary
clone_end:T7
site:EcoRI
end_sequence:BH314503"
  71380..73117
    /note="wgs contig"
complement(242103..242812)
    /note="clone_boundary
clone_end:Sp6
site:EcoRI
end_sequence:BH314504"
  245580..247119
    /note="wgs end_extension
clone_end:Sp6"

Query Match      31.4%; Score 714.803; DB 1; Length 247119;
Best Local Similarity 66.9%; Pred. No. 0.15;
Matches 1379; Conservative 0; Mismatches 537; Indels 144; Gaps 27;

QY 277 TCTCTCCAAACACTTCTATTCTTGATTCATCTTCATCTGGCTCACTTTTAAGTCAAGTGTG 336
Db 39023 TCTCATTCATTCATCAAGGTTTTTAATTTCTGTCCTGACCCACTTTTCATTCAGCAGAG 38964

QY 337 AGTTGTTGGTTTCCATAAGTTTGTAAAGTTTCTGTTTCTGTTCTGTTGTTGTTAT 396
Db 38963 A---GCTCAGCTTCTGTGAGTTGTGAAGCTTCTGTTGTTCC---TGTTGTTATCGATCT 38910

QY 397 CTAGATTTAGCTGTGGTGCATAGGACATAGAGTATTAATTTCAATTTGCTTTTATC 456
Db 38909 CTAACCTTTAATCCACGGTGTCTGATAGGATTCAGTACATTAATCCAAAGTTTCTTGATC 38850

QY 457 TGTCGAGACTTGCTTTGTTTGAATATGATTCATCAATTTTGGAGAGTTTCATAGGGTGT 516
Db 38849 TGTTGAGACTTG-----TGCTCTAAATTTGTTGTAATTTTCCAAAAGACCTATGACATGCA 38794

```

```

QY 517 GACAAGAAGGTACAGCTTTTGTGTTTGGTGAATAGTCTGTAAATATCT-CTAGGTCCA 575
Db 38793 GAAAAGAAGGTATATTTCTTTGTGTTTGGGTGAACGTTCTGTAAATATCTGTAGTTCCA 38734

QY 576 CTTGGTTTATGACATCAGTTAGCTCCAGCATTTTCTCTGTTTCTGTTTCTGTTGATGAC 635
Db 38733 TTTGGTTTGTGATGTCTGCTCTCTCAGCATTTTCTG-TTCAGCTTTTGTCTGGAATGAC 38675

QY 636 CTAACCTGTTGGAGAGAAATGGGTATTGAAGTAGCCCACTATCTGTTGTGAGG-TCAATA 694
Db 38674 CTATCTACTTTGTGAGAAAAGAGATTGAAGCTTACTGCCATCATATGAGGATCAATA 38615

QY 695 TGTGATTTTAGCTGTAGCTGTGCTTTTATGAACTTGGTGCACATTTGTTTGTGTGCA 754
Db 38614 TGTATTTAACCTGAAGTGGTGTTCCTTTTACAACTTGGGTGCCCTTATGTTGGGCA 38555

QY 755 TAGACATTAAGATTCGAATGTCCTCTTGGTGA-TTTTCCTTTTGTAGCTATGTAGTAT 813
Db 38554 TAGATGCCAAGAAATGAAATGTCATCTTGGTGAATTTTCCITGATGATGTAGTGG 38495

QY 814 TCTTCCCAATCTCATCTGCTTAGTTTGGTTTAAAGTCT-ATTAGTCAGATTAATAAG 872
Db 38494 TCTTTCCTATCTC-----TTGTTTGGTTTGAAGTCTGTTTGTAGTAGTAAAGT 38442

QY 873 ACTGTATCGGCTTGCTTCTTAGGGCCATTTGCTTAGAATATCTTTTCCA--TCTTTTAC 930
Db 38441 GCTACACCACCTGGCTTCTTAGGGTCATTTATCTGAAATATCTTTTCCAATCTCTTTAC 38382

QY 931 TCTAAGTGTATGTCATCC--ATGGTAGTGTCTCTTTTGGATGCACAGTAGATGG 987
Db 38381 CCTGAGGCAATGTCATCTCTGATGTCAGGTATATTTATTGGATGCACAGAGATGG 38322

QY 988 ATCTTTTTCATATCCATTTCTGTTACCAGCATCTTTTCTAGAGAAATTAAGATCAT 1047
Db 38321 ATCTATTTCCAGGCGCATCTGTTAGTCTAATGTTT-----AATTTGGTAAATA 38273

QY 1048 GAGTCATTTGATGTTGAGAAATTAATGAGAGAGTGTGTTGGATCTTCTTTATCTTGAC 1107
Db 38272 GAGACCATGGACTTGAGAGATACCAAGGACTAATGATTTGATTCCTCTTACTTCGTTTC 38213

QY 1108 TTGTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1167
Db 38212 TTGTTGTTATGTTGTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 38153

QY 1168 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1227
Db 38152 TATGTATGTATGTGT-----TTCCTTTTGTATTTGCT 38120

QY 1228 GGCTCGGAATTTATTTATTTATTTATTTTCTTGAATGTGGTAAACATCTTTAGATTAAG 1287
Db 38119 GTTCTCAAACTATTTGTTTCTGTTTTCATGGGTATAGTCAACCTCCTTAGGTGAAG 38060

QY 1288 TTTTCTCCTAG--CCTCTTTTAGTCTGCATTTGAAGATAGATATCTTTTACATCTGAT 1345
Db 38059 TGTTCATCTAGAACCTTCTGTAGGCTGGAATTTGTAGACAGATATTAATTTAAATTTGGC 38000

QY 1346 TTTATCTTAGAATGCTTCTTCTTCCAACTATTTGTGACAGAAAGTTTTTCTTAAGTGAG 1405
Db 37999 TTTATCATAGAAATAAATC---TTCTCCATCTATGGTGAATGAATGTTTC-CTGAGTAGAG 37944

QY 1406 TAGTCTGGCTCAGCATCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGCGAGGCTTCT 1465
Db 37943 TAGTCTGGTGGCAACTGTGGTTTCTTGGTATCTGATGAACATCTGCTCAGGTCTCTCT 37884

QY 1466 TACATTTTGAGTTTCTATTTGGAAGAGTCAGGTGTAATTTCTAATACATCTGCTTTATAG 1525
Db 37883 GGCTATTAGAGTCCCACT-GAGAAAGTCATTTGTAATTTTAGTAGGTCTGCTCCTTTATAG 37825

QY 1526 TTAATTTGCTTTTTCCTTCGATCTTTAATATTTCTTTCTTCTTCTTCTTCTTCTTCTTCT 1585
Db 37824 TTGCTTGG--CCCTTCCCTTCGAACTTTTAATATTTCTTCTCTGTCTGTATATTTAGT 37766

```


	* This record will be updated with the finished sequence	
	* as soon as it is available and the accession number will	
	* be preserved.	
	*	
	1 229513: contig of 229513 bp in length	
	229514 229613: gap of unknown length	
	229614 231050: contig of 1437 bp in length	
	231051 231150: gap of unknown length	
	231151 233002: contig of 1852 bp in length	
	233003 233102: gap of unknown length	
	233003 235196: contig of 2094 bp in length.	
FEATURES	Location/Qualifiers	
source	1..235196	
	/organism="Rattus norvegicus"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:10116"	
	/clone="CH230-31M10"	
misc_feature	1..1666	
	/note="wgs_contig"	
misc_feature	168530..169531	
	/note="wgs_contig"	
misc_feature	219147..220600	
	/note="wgs_contig"	
misc_feature	228234..229513	
	/note="wgs_contig"	
	Query Match 31.3%; Score 713.403; DB 1; Length 235196;	
	Best Local Similarity 68.8%; Pred.No. 0.14;	
	Matches 1349; Conservative 0; Mismatches 491; Indels 122; Gaps 28;	
QY	295 ATTCTGTGATTTCTACTTGGCCCAATTTTAACAGTAGTCAGTGTTGGTTTTGCCATA	354
Db	ATTTATCTTTCTGCTTGTGACACATTTTAAAGTCAGTGAAGTTGTTTCAGTTTTCCATG	40649
QY	355 AGTTTGTAAGTTTCTGTTGTTTCTGTTGTTGTTGTTATCTAGATTTAAGCTGTGGT	414
Db	40650 CAITTAAGAATTCCTGTTGTTTCTGTTG-----TAATATCTAGCTTTAATCTGTAGT	40702
QY	415 GGTCAGATGGACATAGAGTATTAATTCATTTGCTCTTTATCTGTCGAGACTTGCTTTGT	474
Db	40703 GGTCTCATGAGTACAGAAGCTATTTCAACCTTCTATCATTGAAGACTTGCTTTGT	40762
QY	475 TTTGAATATGTAATCAAATTTTGGAG--AGTTTCATAGGTCCTGCACAAGAGGTACAGT	532
Db	40763 GTACAAGTATGGTCAAATTTTGGAGAAAGTTCTATAGGTCACAGAGGGAGGTATATT	40822
QY	533 C-TTTCGTTTTCGGTGAATAGTCCTGTAATATCT-CTAGTCCACATTTGGTTTATGACAT	590
Db	40823 CTTTTGGTTTGGGTAAAATGGTCTGTAATTTCTGTTAGGTCATTTGGCTTATAAAT	40882
QY	591 CAGTTAGCTCCAGCATTTCTCTCTTTCGTTTTTGTGAGATGACCTAACCTGTTGGAGAG	650
Db	40883 CTGTAGGCCCCAGTATTTCTCTGTTAG-GTTTGTATCGATGACCTGTCCTGGTGTGAG	40941
QY	651 AATGGGGTATTGAAGTAGCCCATATCTGTGTGAGGTCATATGTGATTTTAGCTGTA	710
Db	40942 AGTGAGCTATGAAAGTCTCCCAGCATCAATGTGTGAGSGTCAATATGTGATTTAGCTCTA	41001
QY	711 GCTGTGCTTTTATGAACTGGGTG----ACATTGTGTTGGTGCATAGACATTAAGA	766
Db	41002 GTGGTCTTTCTTTCCAAATTTGGGTGGTACCCTTGTTGGGAATACAGATGTTAGA	41061
QY	767 ATTGCAATGCTCTTGGTGA--TTTTCCTTTGAGCCTATGTAGTATCTTCCCAATC	824
Db	41062 AGTGAATGGCATTTTGGTGGAAATTTGTCTTTGATGAATATCTACTGTCTT-CCTATT	41120
QY	825 TCATCTGCTTAGTTTTGGGTTTTAAAGTCTA-TTAGTTCAGATATATAAAATGACTGTATCGGC	883
Db	41121 TCTTTTGATGGTTTTTGGTTTGATGTCATTTTGTGAAGACATAAATGGCTACACCAGC	41180
QY	884 TTGCTTCTTAGGCCCAATTTGCTTAGAATATC-TTTTCCAATCCTTTTATCTTAAGGTGATG	942
Db	41181 TTGCTTCTTAGGTCATTTTGGTGGAAATCTTTTCTAGCTCTTTTACCTTGAGATAATG	41240

```

|||||
42260 TTGCTTCTATTCATTTTCAGGCTTGAATGGTTTATTTGTTTCATCCACTGTTGT 42319
QY TTTACATTTTCATAGTTTCTTTTAATGATTTATTCATTTCTCTTCAAGGACCTTTTAT 2104
Db TTTCTATTTCTTGGCTTTCTTTTAAGGATTTACTAATTT--TGCCCAATGTTGTTTA 42377
QY GAAATTCATAAATGATGTTTAAGTCTCTGCTTGTCTTCAGTATGTTGCAATTCAG 2164
Db TGTCTTCTCGAATTTCTTTGAATTCATTTGTTTATTCATTTCTCTATTAATGTTATAT 42437
QY GGCCTATTTGTAATAGGTTTATAGCAGGACATATGTCCTGG 2206
Db AGTTGGTTCTCCTGCTGCTCAGCTATGTTGGAGTATTCAGG 42479

RESULT 23
AC128004
LOCUS Rattus norvegicus clone CH230-262L8, *** SEQUENCING IN PROGRESS
DEFINITION Rattus norvegicus clone CH230-262L8, *** SEQUENCING IN PROGRESS
ACCESSION AC128004
VERSION AC128004.3 GI:25139641
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 247726)
Munzy D.Marie., Metzker M.Lee., Abramzon S., Adams C., Alder J.,
Allen C., Allen H., Albrooks S., Amin A., Anguiano D.,
Anyalebechi V., Aoyagi A., Ayodeji M., Baca E., Baden H.,
Baldwin D., Bandaranaike D., Barber M., Barnstead M., Benahmed F.,
Biswal N., Blair J., Blankenburg K., Blyth P., Brown M.,
Bryant N., Buhay C., Burch P., Burrell K., Calderon E.,
Cadenas V., Carter K., Cavazos I., Cesar H., Center A.,
Chacko J., Chavez D., Chen G., Chen R., Chen Y., Chen Z., Chu J.,
Cleveland C., Cockrell R., Cox C., Coyle M., Cree A., D'Souza L.,
Davila M.B., Davis C., Davy-Carroil L., De Anda C., Dederich D.,
Delgado O., Denison S., Deramo C., Ding Y., Dinu H., Divya K.,
Egan A., Escotto M., Eugene C., Evans C.A., Falls T., Fan G.,
Fernandez S., Finley M., Flagg N., Forbes L., Foster M., Foster P.,
Fraser C.M., Gabisi A., Ganta R., Garcia A., Garner T., Garza M.,
Gebregorgis E., Geer K., Gill R., Grady M., Guerra W., Guvarra W.,
Gunaratne P., Haaland M., Hamill C., Hamilton C., Hamilton K.,
Harvey Y., Havlak P., Hawes A., Henderson N., Hernandez J.,
Hernandez R., Hines S., Hladun S.L., Hodgson A., Hogues M.,
Hollins B., Howells S., Hulyk S., Hume J., Idlebird D., Jackson A.,
Jackson L., Jacob L., Jiang H., Johnson B., Johnson R., Jolivet A.,
Karpachy S., Kelly S., Kelly S., Khan Z., King L., Kovar C.,
Kowis C., Kraft C.L., Lebow H., Levan J., Lewis L., Li Z., Liu J.,
Liu J., Liu W., Liu Y., London P., Longacre S., Lopez J.,
Lorensuhea L., Louisedge H., Lozado R.J., Lu X., Ma J.,
Maheshwari M., Mahindartne M., Mahmoud M., Malloy K., Mangum A.,
Mangum B., Mapua P., Martin K., Martin R., Martinez E.,
Mathiney S., McLeod M.P., McNeill T.Z., Meenen E.,
Milosavljevic A., Miner G., Minja E., Montemayor J., Moore S.,
Morgan M., Morris K., Morris S., Munidasa M., Murphy M., Nair L.,
Nankervis C., Neal D., Newton N., Nguyen N., Norris S., Parks K.,
Nwokedeme O., Okwuonu G., Olarnpunsagoon A., Pal S., Parks K.,
Pasternak S., Paul H., Perez A., Perez D., Pfannkoch C.,
Plopper F., Poindexter A., Popovic D., Primus E., Pu L.-L.,
Puazo M., Quirroz J., Rachlin E., Reeves K., Regier M.A., Reigh R.,
Reilly B., Reilly M., Reuter M., Richards S., Riggs F.,
Rivas C., Rodkey T., Rojas A., Rose M., Rose R., Ruiz S.J.,
Sanders W., Savery G., Scherer S., Scott G., Shatsman S., Shen H.,
Shetty J., Shvartsbeyn A., Sisson I., Sitter C.D., Smajs D.,
Sneed A., Sodergren E., Song X.-Z., Sorelle R., Sosa J.,
Steimle M., Strong R., Sutton A., Svatek A., Tabor P., Taylor C.,
Taylor T., Thomas N., Thomas S., Tingey A., Trejos Z., Usmani K.,
Valas R., Vera V., Villasana D., Waldron L., Walker B., Wang J.,

```

Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstein, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 247726)
Worley, K.C.

Submitted (19-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 247726)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 20, 2002 this sequence version replaced gi:23665065.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size, and separated may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KAJE
Center clone name: CH230-262L8
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 193410 bases at least Q40
Consensus quality: 198067 bases at least Q30
Consensus quality: 201676 bases at least Q20
Estimated insert size: 196651; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 245348: contig of 245348 bp in length
2 245349: gap of unknown length
3 245449: contig of 1110 bp in length
4 246559: gap of unknown length
5 246558: gap of unknown length
6 247266: contig of 1068 bp in length.

Location/Qualifiers
1. 247726
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-262L8"
1. 1135
/note="wgs end extension
clone_end:T7"

FEATURES
source
misc_feature

RESULT 24
 AC095157/c
 LOCUS
 DEFINITION
 AC095157 245174 bp DNA linear HTG 09-MAY-2003
 Rattus norvegicus clone CH230-8N16, WORKING DRAFT SEQUENCE, 3
 unordered pieces.
 AC095157
 AC095157.6 GI:30467202
 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 245174)
 Muzny,D,Marie., Metzker,M, Lee., Abramzon,S., Adams,C., Alder,J.,
 Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
 Anyalebechi,V., Aoyagi,A., Aycoffe,M., Baca,E., Baden,H.,
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
 Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
 Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
 Cardenas,V., Carter,K., Cavazos,I., Cesar,H., Center,A.,
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
 Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
 Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
 Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
 Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
 Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
 Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
 Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
 Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,M.,
 Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
 Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
 Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogue,M.,
 Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
 Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jullivet,A.,
 Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
 Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
 Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
 Lorensuhewa,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J.,
 Maheshwari,K., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
 Mangum,B., Mapus,P., Martin,K., Martin,R., Martinez,E.,
 Mahoney,S., McLeod,M.P., McNeill,T.Z., Neenen,E.,
 Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
 Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
 Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
 Nwaokeleneh,O., Okunolu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
 Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
 Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
 Puzos,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
 Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
 Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
 Sanders,W., Savary,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
 Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
 Sneed,A., Sodergren,E., Song,X.-Z., Sorrelle,R., Sosa,J.,
 Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
 Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
 Valas,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J.,
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
 Williams,G., Willson,R., Wleczyk,R., Woodden,H., Worley,K.,
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
 Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Zhan,D., von
 Niederhauser,A., Weis,R., Smith,D.R., Holt,R.A., Smith,H.O.,
 Weinstock,G. and Gibbs,R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 245174)
 Worley,K.C.
 Direct Submission
 Submitted (16-SEP-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 245174)

AUTHORS
 JOURNAL
 TITLE
 REFERENCE
 AC095157/c
 LOCUS
 DEFINITION
 AC095157 245174 bp DNA linear HTG 09-MAY-2003
 Rattus norvegicus clone CH230-8N16, WORKING DRAFT SEQUENCE, 3
 unordered pieces.
 AC095157
 AC095157.6 GI:30467202
 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 245174)
 Muzny,D,Marie., Metzker,M, Lee., Abramzon,S., Adams,C., Alder,J.,
 Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
 Anyalebechi,V., Aoyagi,A., Aycoffe,M., Baca,E., Baden,H.,
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
 Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
 Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
 Cardenas,V., Carter,K., Cavazos,I., Cesar,H., Center,A.,
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
 Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
 Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
 Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
 Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
 Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
 Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
 Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
 Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,M.,
 Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
 Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
 Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogue,M.,
 Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
 Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jullivet,A.,
 Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
 Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
 Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
 Lorensuhewa,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J.,
 Maheshwari,K., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
 Mangum,B., Mapus,P., Martin,K., Martin,R., Martinez,E.,
 Mahoney,S., McLeod,M.P., McNeill,T.Z., Neenen,E.,
 Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
 Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
 Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
 Nwaokeleneh,O., Okunolu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
 Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
 Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
 Puzos,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
 Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
 Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
 Sanders,W., Savary,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
 Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
 Sneed,A., Sodergren,E., Song,X.-Z., Sorrelle,R., Sosa,J.,
 Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
 Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
 Valas,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J.,
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
 Williams,G., Willson,R., Wleczyk,R., Woodden,H., Worley,K.,
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
 Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Zhan,D., von
 Niederhauser,A., Weis,R., Smith,D.R., Holt,R.A., Smith,H.O.,
 Weinstock,G. and Gibbs,R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 245174)
 Worley,K.C.
 Direct Submission
 Submitted (16-SEP-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 245174)

Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (09-MAY-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On May 9, 2003 this sequence version replaced gi:24817806.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GCUR
 Center Clone name: CH230-8N16
 ----- Summary Statistics
 Assembly program: Atlas;
 Consensus quality: 225865 bases at least Q40
 Consensus quality: 227968 bases at least Q30
 Consensus quality: 229513 bases at least Q20
 Estimated insert size: 237268; sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 241598: contig of 241598 bp in length
 * 241599 241698: gap of unknown length
 * 241699 242839: contig of 1141 bp in length
 * 242840 242939: gap of unknown length
 * 242940 245174: contig of 2235 bp in length.
 Location/Qualifiers
 1..245174
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-8N16"
 2104..3008
 /note="clone boundary
 clone_end:T7
 site:EcoRI
 end sequence:BH347120"
 2142..2962
 /note="clone boundary
 clone_end:T7
 site:EcoRI
 end sequence:BH347120"
 239787..241598
 /note="wgs end_extension
 clone_end:T7"

FEATURES
 source
 misc_feature
 misc_feature
 misc_feature
 Query Match 31.2%; Score 710.102; DB 1; Length 245174;
 Best Local Similarity 69.3%; Pred. No. 0.15;
 Matches 1307; Conservative 0; Mismatches 479; Indels 99; Gaps 26;

QY 435 ATTATTTCAATGTCCTTTATCTCTGAGACTGCTTTGTTTGAATATGTTCAATT 494
Db 63549 AATTTTTCATCTCTTTGTAATCTCTAAGACTTCTGTTGTCGTAATATGTAATCAAT 63490
QY 495 TTGAG--AGTTTCATAGGTCGTGACAAGAGTACAGTC-TTTGTGTTTGGTGAAT 551
Db 63489 TTGAGAAAGTTCCATGAGATGCTAAGAAGAGTATATCTTTTGTGTTGGTGAAT 63430
QY 552 AGTCGTAAATATCTCTAGGTCCT--TGGTTATGACATCAGTTAGTCCAGCATTT 608
Db 63429 GTTCTGTACATATCTGTTGGTTCATCTGGTGTATATAGCATTTGGTTAGTGTAGCATTT 63370
QY 609 CTCTGTTTCTGTTTCTGATGAGTAACTAAGTCTGGAGAGTGGGTATTAAGTAG 668
Db 63369 CTCT----AGTTTATCTAGATGCCCTGTCTATTAGCAAGAGTAGGGTATTGAAGTCT 63315
QY 669 CCCACTATCTGTGTG-AGTCAATATGTAATTTAGCTGTAGCTGTGCTTTTATG 727
Db 63314 CCCACTGTCAAGATGTGAAGTCAATGTGTAATTAAGCTAAAGTAGTGTATGCTTTATG 63255
QY 728 AACTTGGTGACATGTGTTGGTGGCATAGACATTAAGAAATGCAATGCTCTCTGGTGG 787
Db 63254 AACTTCAGTACCTTGTGTTGGTGGCATAGATGTTAAATCAAAATGTCATCTGGTGG 63195
QY 788 A-TTTTCTTTGATGCTATGATATCTTCCAAATCTCAATGCTAGTGTGTTGGT 846
Db 63194 ATTTTCTCTTGACAAGTATAAAGTATCTTCCCTGCTCTCTGTTAGTGTGTTGTTG 63135
QY 847 AAGTCTATTAGTCAGATTAATAATGACTGTATCGGCTTCTCTTAGGGCCATTGCTT 906
Db 63134 ATGTAT-TTTGTGACAAACAAATGGTATACAGCTTACTTANTATACCCATTGCTT 63076
QY 907 AGAATATCTTTTCC--ATCCTTTTACTCTAAGGTGATGCTATPCCATGGTAGTGTGCT 963
Db 63075 AGAGTACCTTCTTCCAACTTTATACCTTACATAATGGTATCCCTTGATTTGAAGTG 63016
QY 964 TTTTGTGATGACAGTAGATGATCTGTGTTTCAATCCATCTCTTACCCAGTATCT 1023
Db 63015 TATTCTTGGATGACAGGAGTAAATCCTGTTTATACATCCATCTCTTAACTGTTAT 62956
QY 1024 TTTTCTAGAGAAATTAAGATCATTGAGTCATGATGTTGAGAAATTAATCAATGACAGTGT 1083
Db 62955 TTTAATTTGGTCAATTGAGA-----CCACTGATTTGGAGATATCAATAGAATGT 62904
QY 1084 TTGAGATCTGTATCTTCACTGTTGAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1143
Db 62903 TTTCTGATCTGTGATTTGGTGTG-GTATGATGATGTTGGTGTGTGTGTGTGTGTGTGT 62845
QY 1144 TGT 1203
Db 62844 GGT 62788
QY 1204 GTTCTCTCCCTCTTTGATTTTGGCTGGAATATTTATTT--ATTCAATTTTCTTGA 1261
Db 62787 GTT-----TTCCGTGCTGAGATTAATTTATCCCTTTTTCACCTTTG 62744
QY 1262 ATGTTGGTAACTCTTTAGATGAAGTTTCTCTA---GCCTTCTTTAGGTCGATTT 1318
Db 62743 GTGTGGGAGGCTCTTTAGACAGAGTTAACTTATAGTAATTTAGTAGAGCTTGATTT 62684
QY 1319 TGAAGATAGATATCTTTTACATCTGATTTTATCTTTAGAAATCTTTCTTCTTCAACTAT 1378
Db 62684 TATATATAAATTTGCTGTGATTTAGTTTATCATATAAATGCTGTGTCTTTCTTCTATAT 62624
QY 1379 TGTGACAGAAAGTTTTCTTAAGTCAGTAGTCTTGGCTGACATCTGTAGTCTCTTGGAGT 1438
Db 62623 TGTGATTTGAAAG---TGTGGGTATAAATGCAATGCAAGGCTGGCATCTGTGGTCTTTAGCCT 62567
QY 1439 CTGTAGCACATCTGGAGGCTCTTACATTTTGTAGTTTCTTATTTGAAAGTCAAGTG 1498
Db 62566 CTGTAGCACATTTGCTAGGCTCTTTTGTCTTTTAGAGTCTTCCA-TGAAAGTCAAGTG 62508
QY 1499 TAAATCTAATACATCTGCTTTTATATGTTAATGTTCTTTTTCCTTTTCCCTTGTCACTTTTAA 1558

Db 62507 TAGTCTGATAGTCTGCTTTGTATGTACTTGTGTTTCTCCCTGACGCTTTTAA 62448
QY 1559 ATCTTTC-TTGTGCTATACCTTTAGTAGTATTAATATGCACTGTGGGAGTTCCT 1617
Db 62447 ATCTTTCCTTTTCTGTACATTCAGTGTTTGATTAATATGTCGAAGAGACTTTCT 62388
QY 1618 TTTCCGGTCCAACTATTTTGGTGTGTTGATGCTTCTTGTACCTTGATAGGCACTCTCTT 1677
Db 62387 TTTCTGGCTCAATCTATTTGGTATTTCTGTATGCTGATGCT-----TACCTTTAGGCACTCTCTT 62334
QY 1678 CTCAAGGTTAGAAATTTTCTTTTGTGTTTCTTTGAAATATTTTCCCTGCTTTGAC 1737
Db 62333 CTTTATGTTAGAAATTTTCTTCTATGATTTGCTGAAATATTTTCTGTGCTCTTGAC 62274
QY 1738 CTGCTTCTTCC-CTTCTCTCTATTCC-----TTTGGTGTGTCATAGTGTCT 1783
Db 62273 CTGGATCCCCCTCTCTCTTCTTATCTTATTAATCTTTTATAGTGTCT 62214
QY 1784 TCTGGCTTCTCGATGTTTATGCTGGAATTTTATAGACTTTAACTTTCTTTGACCAA 1843
Db 62213 CCATATTTCTCGACATTTTATACAGGAGTCTTTAGAGTAAACATTTCTTTAACCAA 62154
QY 1844 GGTATCCATTTCTTATCTTCTCTCTCTGCTGCTGAGATCTCTCTATCTCTGAT 1903
Db 62153 GGCATTCATTTCTTATGCTCTTCAATGCTCGAGAT--TCTCTCATCTTTTGTGT 62096
QY 1904 TCTGTGAGTGAGCTGTCTCTGAGTCTCTGCTGGTCTTAAATTTTCTCAATTTCCAGA 1963
Db 62095 TCTGT-----TGCCCTGAGTCTTTTCCGGTGTCTTATTTCTCAATTTCAAGA 62045
QY 1964 TTTCTTCAAGTTGGTGTGTTTATTAATCTATTTCCACTTTCAGTCTCGAATGTT 2023
Db 62044 TTTCCCTCAGTTGGGTTTCTTTATAGTTTATTTTATACATTCAGATCTTTGAATGCT 61985
QY 2024 TTACTCATTTTCTCCAGTATTTACATTTTCTAGGTTTCTTTAATGATTTATTCATT 2083
Db 61984 TTATCTA-TTCTTCTCCACTGTTGTTTTCATA-GATCTTAAAGATTTATTTAT 61927
QY 2084 TCTCTTCAAGGACCTTTTATGAATTCATAAATGATGTTAAGTCTTGCCTGTGTGT 2143
Db 61926 TCTCTTTTAAAGATCTCTATCATTTGTCATAAAGGCTGTTTGAAGTCTTATTTGTGT 61867
QY 2144 TCAGTATGTTGCATCTCAGGGCT-----TATGTAATAGGTTTACAGGAC 2194
Db 61866 TCAGTATGCTGGAATATTCAGGGCTGCTGTGGTAGGGTGGGGGTTCTAGTGAGAC 61807
QY 2195 ATATTCTCTGTTGTTTATGTTCTGTTTGTCTTTGCTTTGGCATATAGCGCTGAGTTGG 2254
Db 61806 ATATAGTCTGACTGTTATAGTTGTTGATCC--TAGCATCTGGCATCTGGGTTGG 61749
QY 2255 GATGATTGAATCTAGGTGCTGAT 2279
Db 61748 AATGATTGAATTTTCAAGTGTTCAT 61724

RESULT 25

AC113514
LOCUS 167223 bp DNA linear HTG 21-OCT-2002
DEFINITION Mus musculus clone RP23-475D11, WORKING DRAFT SEQUENCE, 9 ordered
pieces.

AC113514
VERSION AC113514.4 GI:24182231
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 167223)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Birren, B., Nusbaum, C. and Lander, E.

TITLE Mus musculus, clone RP23-475D11

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 167223)

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B., Brown,A., Canarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cook,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Girde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A., Karatas,A., Kells,C., LaRoque,K., Lamazares,R., Landers,T., Lehoczyk,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Mathews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Meneus,L., Minova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Roman,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Royman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembeck,L., Zimmer,A. and Zody,M.

TITLE

Direct Submission

JOURNAL

Submitted (01-MAR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

3 (bases 1 to 167223)

AUTHORS

Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Canarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Kanat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Matthews,C., McCarthy,M., Meldrum,J., Meneus,L., Minova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembeck,L., Zimmer,A. and Zody,M.

TITLE

Direct Submission

JOURNAL

Submitted (21-OCT-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On Oct 21, 2002 this sequence version replaced gi:21326539.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

1 10572: contig of 10572 bp in length
10573 10672: gap of 100 bp
10673 12088: contig of 1416 bp in length
12088 12188: gap of 100 bp
12188 13362: contig of 1174 bp in length
13362 13463: gap of 100 bp
13463 12741: contig of 9279 bp in length
12741 22841: gap of 100 bp
22841 27991: contig of 5150 bp in length
27991 28091: gap of 100 bp
28091 40867: contig of 12776 bp in length
40867 52499: contig of 11532 bp in length
52499 52500: gap of 100 bp
52500 154378: contig of 101779 bp in length
154378 154479: gap of 100 bp
154479 167223: contig of 12745 bp in length.

FEATURES

source

1..167223
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-475D11"
/clone_lib="RPCI-23 Female Mouse BAC"

misc_feature

1..10572
/note="assembly_fragment
clone end:SP6
vector side:left"

misc_feature

10673..12088
/note="assembly_fragment"
12189..13362
/note="assembly_fragment"

misc_feature

13463..22741
/note="assembly_fragment"
22842..27991
/note="assembly_fragment"

misc_feature

28092..40867
/note="assembly_fragment"
40968..52499
/note="assembly_fragment"

misc_feature

52600..154378
/note="assembly_fragment"
154479..167223
/note="assembly_fragment"

misc_feature

clone end:17
vector_side:right"

Query Match 31.0%; Score 707.202; DB 1; Length 167223;
Best Local Similarity 67.5%; Pred. No. 0.12;
Matches 1385; Conservative 0; Mismatches 538; Indels 130; Gaps 29;

QY	275	TCCTCTCCAAACACTTCTATTTCTTCTGATTCATCTTGGCTCATTTTAACTCAGTAG	334
Db	71989	TCAGTTTGTAGAAAGTCTTTAAATTTCTTTTCTGTTTGCACATTTT--TTTCAGTAG	72046
QY	335	TCAGTTGTGTTT--GGTTTCCATAAGTTTGTAAAGTTTCTGTTTCTGTTGTGTTGTGT	393
Db	72047	ACAGTTGTTTGTAGTTTTCATGATGTTTGTAAAGCTTCTATTATTTTC--TTTGTGTTGA	72103
QY	394	TATCTAGATTTTAAGCTGTGGTGGTCAGATAGGACATAGATATTATTCATTGCTTTT	453
Db	72104	TATCCATCTTTTAATCCGTGGTGGCTCGAGAAATAAGTAGTAGTTAGTTCAA-TGTTCTTG	72162
QY	454	ATCTGTCCGAGACTTCTGTTTGTGTTTGTAAATATGTTATTCATTGTTG-GAGAGTTTCATAGGG	512
Db	72163	ATCTGTTGAGACTTCTGTTTGTGTTTGTGTTGATGATCAATTTTAGAGAAAGCTCATAGG	72222

* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. Gaps between the contigs

```
QY 513 TGCTGACAAGAAGTACAGT---CTTTGCTTTTGGTGAATAAGTCTGTAAATATCT-CT 568
Db 72223 TATAGAGAAGAACTGTATATCTTTTGTGTTGGGTGAGATGTTCTGTAAAATATATT 72282
QY 569 AGGTCCACTTGGTTTATGACATCAGTTAGCTCCAGCAATTTCTCTGTTTCTGTTTGTG 628
Db 72283 AGGTCCACTTGGTTTATGACATCAGTTAGCTCCAGCAATTTCTCTGTTTCTGTTTGTG 72341
QY 629 AGATGACCTAACTGTTGGAGAGAAATGGGATTAATGAAGTAGCCCACTATCTGTGTGTGAGG 688
Db 72342 AGATGACTTGTCTATTTGTAATAGTGTGTAATGAAGTCTCCACCGTCACTGTGTGTGAGG 72401
QY 689 TCAATATGTGAATTTTAGCTGTAGCTGTCTGTTTATGAAGACTTGGTGACATTTGTGTTT 748
Db 72402 GTA-----AGPAGTGTTTATTTTACAAAGGTGGGTG----- 72432
QY 749 GGTCATAGACATTAAGAAATGCAATGCTCTCTGTGTGGA-TTTTCCCTTTGATGCCCTATG 807
Db 72433 ---GCATAGATGTTAATAAGAAATATCATCTTGGTGGATTTTCCCTTTGATGAGTATG 72489
QY 808 TAGTATTTCCCAATCTCATCTGCTTAGTTAGTTTGGGTTTAAAGTCTAT-----TAGT 858
Db 72490 TTGTGTCCTTCCCTGCTCTCTTAATTAGTTTGGTCTGAAGTCTGTTTATTTGATATC 72549
QY 859 CAGATATTAAGATGACTGTATCGCTTCTCTTAGGGCCATTTGCTTAGAATATC----- 914
Db 72550 AAAATAGCTACAACTTAGGACCTTTTCTTAGGTCATTTGTTTAGAAATATCTTT 72609
QY 915 -TTTTCCATCTTTACTTAAGTGATGTCATC---CATGGPAGTGTCTCTTTTGTG 970
Db 72610 TTTTGTGAACCTTTACCTTAGTAAATTTCTATCTTTGATTTGAAGTATATTTCTTGG 72669
QY 971 ATGCAGCAGTAGGATGATCTTTTTCATATCTGTTACCCAGTATCTTTTCTTA 1030
Db 72670 ATGCAGCAAAAGGATGATCTGTTTGGCAATCTTTCTATTAGACTGTGCTTTTATT 72729
QY 1031 GAGAAATTAAGATCATTTGAGTCATGATGTTGAGAAATATCAATGAGCAGTGTGTTGGA 1090
Db 72730 GGGGATTTAGACCAATGA-CCATTTGACATGAGAGATCAGTCCCAATAATGTTGA 72788
QY 1091 TTTCTGTTATCTTGCACTGTGAAGTGTGTGTG-TGTGTGTGTGTGTGTGTGTG 1149
Db 72789 TTTCTGTTGTTTGTGTTTATTTATTTGTTGGTGTGATGTTGTGTACACATGTTCA 72848
QY 1150 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1209
Db 72849 TTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 72905
QY 1210 TCCCTCTTTTGATTTTGGCTCGAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTT 1269
Db 72906 -----TTAGATTGTTTATTTTCCATGTTTGTGTTTATAGTT 72943
QY 1270 AACATCTTTAGATGAAGTTTTCCTCTAG--CCTTCTTTAGGTCGCAATTTAGAGTAG 1327
Db 72944 TTTCTCTTGTGTTGAGTTTTCCTTCCAGACCCCTCTGCAGGACTGGAATTTGATAGTAG 73003
QY 1328 ATATTTCTTACATCTGAT-TTTATCTTGAATGCTTTCTTTCTTCCCAACTATTGTGACAG 1386
Db 73004 ATCAATATTTGTTTAAAGTGTTTTAAAGGAATATCTATTTCTTCCATCTGTAGTACTG 73063
QY 1387 AAAGTTTCTTAAGTGCAGTGTGTGCTGACATCTGTAGTCTCTTGGAGTCTGTAGCA 1446
Db 73064 AGAGTCTTGTCTGGGTATAGTAGTGTGGCTGACATCTGTGTTTCTTATAGTCTGACGA 73123
QY 1447 CATCTGTGCAGGCTCTTCAATTTTGAATTTTCTATTTGAAAGTCAAGGTGTAATTTCTA 1506
Db 73124 CATCTGTCCAGATCTTCTTCTGCTTTTAGT-----CCATGAGAGTCAAGGTGTAATTTCTC 73180
QY 1507 ATACATCTGCCCTTATATGTTAAATGGTCTTTTTTCCCTTCCATCTTTTAAATATCTTTC 1566
Db 73181 ATAGTCTGTCTTTGTATGTTACTGTGTC-TTTTCCCTTGTAGCTTTTAAATATCTTCCC 73239
```

```
QY 1567 TTTGTTCTATACCTTTTAGTGATTGATTTATTATGACATGTGGGAGTTTCTTTCCGGTC 1626
Db 73240 TTTGTTCTGTACATTAATAATTTTGAATTAATTAAGTCTGAGGGGCTTTCTTTTCTGTT 73299
QY 1627 CAATCTATTGTTGTTTGTATGCTTCTGTACCTTGTATAGGCATCTCTTTCTCAAGTT 1686
Db 73300 CAATCTACTAGTGTCTATATGCTTCTGTACCTTTATAGSCATCTCATCTTTAGGTT 73359
QY 1687 AGGAAATTTTCTTTTGGTGTCTTGAANAATATTTTCCCTGCTTTTGGACCTG----- 1740
Db 73360 AGGAAATTTTCTCTATCATTAATTTGAAATAGTTTCTGGGCTTTTGGAGTGAATTC 73419
QY 1741 -CCTTCTTCCCTCTCTCTATTC-----TTTGGTTTTGCAVAGTCTCTCT 1786
Db 73420 TCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 73479
QY 1787 GGTCTCTGATGTTTATGCTCTGATTAATTTAGACTTAACATTTCTTTGACCAAGT 1846
Db 73480 AATTTCTGTTGTTTGTGTC---AGCTTTTATAGATTAACATTTCTTTTGA---GGTGT 73534
QY 1847 ATCCATTTCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1906
Db 73535 ATTTATTTCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 73594
QY 1907 GTCAAGTGGCTTGTCTGAGGTTCTCTGTTGGTCT---TAATTTTTCATTTCCAGA 1963
Db 73595 ATAGGTGAAGCTTGCCTCTGTAGTTACGTTTCAATTTCTAAATTTCTTTTTCAGA 73654
QY 1964 TTTCTCTAGTTGGTGTGTTTATTAATCTATTTCCACTTTCCAGTCTCTGAAATGTT 2023
Db 73655 ATTCCCTCAGTTGGGTTTCTTTATGATCTGTTTCCATTTTCAAGTATTGAACAGTT 73714
QY 2024 TTACTCATTTTCTCTCCAGTATTTACATTTTTCATAGGTTTCTTTAATGGAATTTTCAAT 2083
Db 73715 TTATTTCA-TTTCTTCCACCCCTGTGTTTATAGATTTCTTAAAGCAATTTATTAATC 73773
QY 2084 TCTCTTCAAGCACTTTTATGAATTCATPAAATGATGTTAAGTCTCTGCTGTGCT 2143
Db 73774 TCTCTTTAAGGATCTCTGTATCTTCAATTCAGTGGCTTTAAAGTCTTTTCTTATAT 73833
QY 2144 TCAGCTATCTTGA---TTCTCAGGGCTATTGTAATAGGTTTGTAGCAGGACATATG 2200
Db 73834 TCAGCTATGTTAAAGGGCTCTGCTGTGATAGAGGCTCTCCATGAGACAAATG 73893
QY 2201 TCTGTTGTTATTTGTTCTGTTTGTCTTTTGGCATATAGCGGCTGAGTTGGGATGAT 2260
Db 73894 CCCCAG-TGTATTAATGTTGTTTTATGTTAGCATCTAGGTATCTGGAGTTGGGATGAT 73952
QY 2261 TGTAAATCTAGGT 2273
Db 73953 TGTAGTCTAGTT 73965
```

RESULT 26

```
AC110100 250188 bp DNA linear HTG 20-NOV-2002
LOCUS Rattus norvegicus clone CH230-212E13, WORKING DRAFT SEQUENCE, 2
DEFINITION
unordered pieces.
AC110100
VERSION AC110100.6 GI:25139572
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 250188)
Kuzny,D.Marie., Vetzler,M.Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebech,V., Ayayi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
```


Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A., Chacko, J., Chavez, D., Chen, Y., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, R., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gregoire, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Guaratone, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hughes, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensu, H., Louised, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokemele, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, B., Primus, E., Pu, L.-L., Puzo, M., Quirroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rokey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, M., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steime, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worthy, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhauser, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Unpublished
2 (bases 1 to 250188)
Worley, K.C.
Direct Submission
Submitted (10-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 250188)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 20, 2002 this sequence version replaced gi:23270205.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GOFV
Center clone name: CH230-212E13
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 227481 bases at least Q40
Consensus quality: 230202 bases at least Q30
Consensus quality: 232076 bases at least Q20
Estimated insert size: 239268; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 33011: contig of 33011 bp in length
* 33012 33111: gap of unknown length
* 33112 250188: contig of 217077 bp in length.
Location/Qualifiers
1..250188
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-212E13"
31752..33011
misc_feature
/note="wgs_contig"
33112..35207
misc_feature
/note="wgs_contig"
Query Match 31.0%; Score 706.902; DB 1; Length 250188;
Best Local Similarity 69.2%; Pred. No. 0.16;
Matches 1305; Conservative 0; Mismatches 481; Indels 99; Gaps 26;
QY 435 ATTATTCATGCTCTTTTATCTGCGAGACTGCTTTGTTGTAATATGATTCATTT 494
Db 48760 AATTTTCAATCTCTTCTGTAAGACTTGTCTGCTGCTGAATGATGATCAAT 48819
QY 495 TTGGAG--AGTTTCATAGGCTGCTCACAAGAGGTACAGTC--TTTGTCTTTGGTGAAT 551
Db 48820 TTGGAGAAAGTTCCATGAGATGCTAAGAGAGGTATATTTCTTTGTTGGTGGAT 48879
QY 552 AGTCTGTAATATCTCTAGTCCACT---TGTTTATGACATCATGTTAGTCCAGCATTT 608
Db 48880 GTTCTGTACATATCTGTTGGGTTTCATCTGGTGTATATAGCATTTGTTAGCTGTAGCATTT 48939
QY 609 CTTCTGTTCTGTTTGTGATGATCCTTAAGTGTGGAGAGATGGGGTATTGAAGTAG 668
Db 48940 CTCT-----AGTTTTATCTAGATGCCCTGCTATTAGCAAGATAGGTATGAGTCT 48994
QY 669 CCACATCTATGTTGTG-AGGTCAATATGTCATTTTAGCTAGCTGCTGTTGTTTATG 727
Db 48995 CCACCTGTGAGGATGGAAGTCAATGTGTCATTAAGCTAAAGTAGTGTATGCTTTATG 49054
QY 728 AACTTGGGTGATGCTGTTTGGTCATAGACATTAAGAATGCAATGCTCTCTTGGTG 787
Db 49055 AACTTCAGTACCTTGTGTTTGGGTCATAGATTTAAATCAAAATGTCATCTTGGTG 49114
QY 788 A-TTTTCTTTGATGCCCTATGATGATTTCTTCCCAATCTCATCTGCTTAGTTTGGGTTT 846
Db 49115 ATTTTTCCTTTGACAAGTATAAGTATTTCTTCCCTGCTCTCTCTGTTAGTTTGGTTG 49174
QY 847 AGTCTATAGTCAGATATTAAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTT 906
Db 49175 ATGTAT-TTTGTGACACAACAAATGGCTATACAGGCTTACTTATTATACCCATTTGCTT 49233

QY 907 AGAATATCTTTTC---ATCCTTTTACTCTAAGGTGATGCTATCCATGCTAGGTTGTCT 963
Db 49234 AGAGTACCTTCTCCAAACACTTTATACCCCTACATAATGCTATCCTTGATGTTGAAGTG 49293
QY 964 TTTTGTGATGACAGTAGAGTATGATCTGTTTTCATATCCATCTGTTACCCAGTATCT 1023
Db 49294 TATTTCTTTGATGACAGGAGTAAATCCTGTTTTCATCCATCTGTTTAACTGTATAT 49353
QY 1024 TTTTCTAGAGAAATTAAGATCAATGAGTCATTTGATGTTGAGAAATTAATCAATGACGAGTGT 1083
Db 49354 TTTAATTTGGGTCATTTGAG-----CCACTGATGTTGGAGATATCAATAAGAAATGT 49405
QY 1084 TTGTGGATCTTTGATTTCTTGCACTGTTGAGAGTGTGTTGTTGTTGTTGTTGTTGTTG 1143
Db 49406 TTTCTGATTTCTGTTGATTTTGGTGGTG-GTAGTGAAGTGGTGGTGGTGGTGGTGGTGGT 49464
QY 1144 TGT 1203
Db 49465 GGTGGT 49521
QY 1204 GTTCTCTCCCTCTTTTGAATTTTGGCTGGAATTAATTAAT--ATTCATATTTTCTTGA 1261
Db 49522 GTT-----TTCTGGTGTGAGATTAATTAATTCCTTTTTCACCTTTG 49565
QY 1262 ATGTGGGTAAATCTTTAGATTTGAAGTTTTCCTCTA--GCCTTCTTTAGGTCGATTT 1318
Db 49566 GTGTGGGAGCCTCTTTAGACGAGAGTTAACTTATTAAGTAATTTCAATGAGCTGATTT 49625
QY 1319 TGAAGATAGATATTTCTTTACATCTGATTTTATCTTGAATGTCCTTTCTTTCTCAACTAT 1378
Db 49626 TATATATAATATTTGCTTGCATTTAGTTTATATATATAAATGTCCTTTCTTTCTCATCTAT 49685
QY 1379 TGTGACAGAAAGTTTCTTAAGTGCAGTAGTCTGGCCTGACATCTGTAGTCTCTTGGAGT 1438
Db 49686 TGTGATTTGAAAG---TGTGGGTATTAACAATGACGGCTGACATCTGTGGTCTCTTAGCCT 49742
QY 1439 CTGTAGACATCTGTGAGGCGCTCTTACATTTTGTAGTTTCTATTTGAAAAGTCAAGTGT 1498
Db 49743 CTGTAGACATCTGTGAGGCGCTCTTCTTTTGTAGTCTTCA-TGAAAAGTCAAGTGT 49801
QY 1499 TAATCTAATAACATCTGCTTTATATGTTAATGCTCTTTTTCCTTTCCCTGTCATCTTTAAT 1558
Db 49802 TAGTCTGATAGTCTGCTTTGTATGTTACTTGTCTTTTCTCTCTGAGCTTTTAAAT 49861
QY 1559 ATTCTTTCT-TTTGTCTATATCTTTTGTAGTATTCATTTATATGACATGCTGGAGTTTCT 1617
Db 49862 ATTCTTTCTTTTCTGATCTCAGTCTGTTGTTGATTTATGTTATGTTGTTGTTGTTGTTG 49921
QY 1618 TTTCCGTTCCAAATCTATTTGTTGTTTGTATGCTTCTTTGTACCTTTGATGCTATCTCTTT 1677
Db 49922 TTTCTGGCTCAATCTATTTGTTATTTCTGTATGC-----TACCTTTCTAGGCTCTCTTT 49975
QY 1678 CTCAGGTTAGAAATTTTCTTTTGTGTTTCTTGAAAATATTTTCCCTGCTTTTGAC 1737
Db 49976 CTTTATGTTAGAAATTTTCTCTATGATTTTCTGAAAATATTTTCTGTGCTTTTGAC 50035
QY 1738 CTGCTCTCTTCC-CCTTCTCTATTTCC-----TTTGGTTTGTGATGATGTC 1783
Db 50036 CTGGATCCCTCTCTCTCTATTTCTTATTTCTTAGATTTAATCTTTTATAGTCC 50095
QY 1784 TCTGGCTCTCTGATGTTTATGCTGATTTATTTTATGATTTAATCAATTTCTTTTGACCA 1843
Db 50096 CCATATTTCTCTGGACATTTTATACAGAGTCTTCTAGAGTAAACATTTCTTTTAAACCA 50155
QY 1844 GGTATCATTTCTTCTATCTTGTCTTCACTGCTGAGATTTCTCTCTCTATCTCTCTGAT 1903
Db 50156 GGCATCATTTCTTCTATGTTGTTCTTCAATGCTGAGAT--TCTCTTCAATCTTTTGT 50213
QY 1904 TCTGTAGTAGGCTTGTCTCTGAGGTTCTGTTGGTCTTAAATTTTCAATTTCCAGA 1963
Db 50214 TCTGT-----TGCCCTCGAGTCTTCTTCCGGTGTGTTAAATTTCTCAATTTCAAGA 50264
QY 1964 TTTCTCTCAGTTTGGGTTTGTGTTTATTAATTTCTATTTCCACTTTTCAGGCTCTGAAATGTT 2023

Db 50265 TTTCCTCAGTTTGGGTTTCCCTTTATTAGTTTATTCTACATTCAGATCTTGAAGTCT 50324
QY 2024 TTACTCATTTTCTCCAGTATTTTACATTTTTCATAGGTTTCTTTAATGATTTATTTACT 2083
Db 50325 TTATTTCA-TTCTTTTCCACTGTTTGTGTTTTCATA-GATTCCTTAAATGATTTATTT 50382
QY 2084 TCTCTTCAAGACCTTTTATGAATTCATAAATGATGTTAAAGTCTCTTGCCTGTGCT 2143
Db 50383 TCTCTTTTAAATCTCTATCATTTGTCTATAAAGGCTGTTTGTAGGTTCTTATTTTGTGTT 50442
QY 2144 TCAGTATGTTGCAATTCACAGGCC-----TATTGTAATAGGTTTATAGCAGGAC 2194
Db 50443 TCAGCTATCTCTGGAATATTCAGGGCTGCTGTGGTAGGTTGGGTTTCTAGTGGAGAC 50502
QY 2195 ATATTGCTGCTGTTTATTTGTTCTGTTTCTGTTTTCCTTTCGATATAGAGCGCTGAGTTGG 2254
Db 50503 ATATAGTCTTGAATTTAGTTTGTGTTTGTATCC--TAGCATCTGGCATCTGGGTTGG 50560
QY 2255 GATGATTTGAATTTCTAGGTCGTGAT 2279
Db 50561 AATGATTTGAATTTTTCAGTGTTCAT 50585

RESULT 27
AC129113 168169 bp DNA linear HTG 19-NOV-2002
LOCUS Rattus norvegicus clone CH230-34207, WORKING DRAFT SEQUENCE.
AC129113
AC129113.4 GI:25072866
HTG: HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 168169)
AUTHORS Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Delgado, O., Denson, S., Deramo, C., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebrgeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hernandez, J.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karnathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kratz, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshew, L., Louised, H., Lozano, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidas, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K.,
Nwokedeme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pflankoch, C.,
Plommer, F., Poindexter, A., Popovic, D., Primus, E., Pu, L. L.,
Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P.,


```
QY 1062 GAGAAATATCAATGAGCAGTGTGTTGTGGAATCTTTGTTATCTTTGACATCTGTGAAGTGTGTG 1121
Db 131984 TAAAGATATCAATGAGCAATGAAATGTTAACTCCAGTTATTTTATGTTG-GTGCTGCTG 132042
QY 1122 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1181
Db 132043 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG- 132096
QY 1182 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1241
Db 132097 -----TTCCCTCTCTTTCAATTTTGTGTGTGTGAATATT 132132
QY 1242 TATTATTCATATTTCTTGAATGTTGGGTAAACATCTTTAGATGGAAGTTTTCCTCTAG-- 1299
Db 132133 TATATCTGTGTGTTTCTTGTGGCAGATGTTAACTCTCTGGGTAGAGTTTTCCTCTAGTA 132192
QY 1300 CCTTCTTTAGCTGTGCAATTTGCAAGATAGATATCTTTACATCTGATTTTATCTTAGAATG 1359
Db 132193 TCTTCTATAGGGCTGGATTTGTGATAAATATGCTCTAAATTTGGTTTGTGTGATGCAATA 132252
QY 1360 TCTTTCTTTCTCCAACTATTTGTGACAGAAAGTTTCTTAAAGTGCAGTAGTCTGGCCT--G 1417
Db 132253 TCTTATTTCTCCATCATGATAATTGAGTGTGTTTGTGCTGATTTAAAGTCTGGCTGAG 132312
QY 1418 ACATCTGTAGTCTCTTGGAGTCTGTAGCAGATCTGTGCAGGGCCTTCTTACATTTTGAAT 1477
Db 132313 ACATCTGTGTCTCTCAGAGTCTGGAAGACATCTGCCAGGCCCTTCTGACTTTTAGAT 132372
QY 1478 TTCTATTGGAAAGTCAAGTGTAAATCTTAAATACATCTGCCCTTTATATGTTAAATTTGTCCT 1537
Db 132373 CTCTGAT-GAGAAAGTGTAGTGTAAATCTTAAAGGCTGCCCTTTATGTTTATTTGG-CAT 132430
QY 1538 TTTTCCCTTGCACTTTTAAATCTTTCTTTCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1597
Db 132431 TTTTCCCTTGCACTTTTAAAT-TCCTTTCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 132489
QY 1598 ATGCACTGTGGGAGTCTTCTTCCGTTCCAACTATTTTGTGTGTTTGTGTGTTGTTGTTCTGT 1657
Db 132490 AAGTGTGTGGAAGATTTCTTTCTGATCCAACTGTTTGTGTGTTTGTGTGTTTGTGTGTTGTT 132549
QY 1658 ACCTGTATGAGCATCTCTTCTCAAGGTAGGAATTTTCTTTTCTTTTGTG-GTTTCTTTGAA 1716
Db 132550 ACAATTTAGGCAGTCTCTCTTTTAGTTAGAGAAATTTCTTCTATGACTTTTGTGTTTAA 132609
QY 1717 AATATTTTCCCTGTTTGTGCTG-----CCTTCTTCCCTTCTTCTTCTTCTTCTTCTTCTT 1766
Db 132610 ACTATCTTCTGGGCTTTTAACTGGGAATCTTCTTCTTCCAACTCCACTTCTTCTTCTTCTT 132669
QY 1767 GGTTTTTCATAGTCTCTCTGGCTTCTGATGTTTATGCTGGATTTATTTAGACTTA 1826
Db 132670 TGGTTTTCTAGTGTCCAGATTTCTGATGTTTGTGTGTCAGAACTTTCTAGATATA 132729
QY 1827 ACATTTCTTTGACCAAGGTATCCAACTTCTTCTATCTTGTCTTCTTCTTCTTCTTCTTCTTCT 1886
Db 132730 ACATTTCTTTGACTATGATGATGATTTCTTCTTATGTAACCTTCTTCTTCTTCTTCTTCT 132789
QY 1887 TCTTCTATCTCTGTGATTTCTGTGAGTGGCTTGTCTGAGGTTCTGTTGGTCTTCTTA 1946
Db 132790 TCTCCATCTCTTGA-----TCTGTGTCTCTGTCTTCTTCTTCTTCTTCTTCTTCTT 132829
QY 1947 ATTTTTCATTTCCAGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2006
Db 132830 AGGTTTTCCATCTCCAGGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 132889
QY 2007 TCAGTCTCTGAAATGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2063
Db 132890 TGAAGTCTGGAACGTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 132949
QY 2064 CTTTAAATGGATTTATTCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2123
Db 132950 CTTTATGGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 133009
```

```
QY 2124 TAAGTCTCTTGTGTTGTCTTCAAGTATGTTTCAATTTCTCAGGCT-----ATTG 2173
Db 133010 TAAGTCTATTTTCTGTGTTTCAAGTGTGTTAAGACATCCAGGCTTGCATCATGAAGAG 133069
QY 2174 TATAGGTTTATAGCAGGACATATGCTCGTGTGTTTCTGTTTCTGTTTCTTCTTCTTCTTCTT 2233
Db 133070 AAGCTGGATTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 133129
QY 2234 C 2234
Db 133130 C 133130

RESULT 28
AL683804
LOCUS
DEFINITION Mouse DNA sequence from clone Rp23-17F2 on chromosome 1, complete
sequence.
ACCESSION AL683804
VERSION AL683804.15 GI:24366485
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 198631)
AUTHORS Skuce, C.
TITLE Direct Submission
JOURNAL Submitted (16-OCT-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Oct 24, 2002 this sequence version replaced gi:24210306.
COMMENT ----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
-----
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; Sw:
SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep Rp23-17F2 is
from the RPCI-23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBAC6.6 Location/Qualifiers
source
1..198631
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="1"
/clone="Rp23-17F2"
/clone_lib="RPCI-23"
```

Query Match 30.9%; Score 704.103; DB 1; Length 198631;
Best Local Similarity 65.1%; Pred. No. 0.14;
Matches 1355; Conservative 0; Mismatches 574; Indels 153; Gaps 23;

Qy	270	TTCCCTCTCTCTTCCAAACACTCTATTTCTTTGTAATTTCTTAATCTTGGCTCATTTTTAACTC	329
Db	186779	TTCACTGAATTTTAGAAAGCTTAAATTTCTTTCTTTATTTCTTCCCTAA---TCATTG	186834
Qy	330	AGTAGTGAGTCTTTGGTTTCCATAAGTTGTGAAGTTTTCTGTTTCTTCTGTTGTGTG	389
Db	186835	AGTAGAGCCATTACAGTTTCTCMAATATGTAGGCTTCTGTGTTCCTGTGTTGTGTG	186894
Qy	390	TTGTATCTAGATTTAAGCTGTGGTGTCAGATAGGA-CATAGAGTATTTATTTCAATTGT	448
Db	186895	AAG---TCCAGCTTTATTCCATGGTGAICTGATAAGATCCATGGGTTTATTTCATCTT	186951
Qy	449	CTTTATCTGTGAGACATCTTTGTTTGAATATGTAATTAATTTTCGAGA---GTTTC	506
Db	186952	CTGTATCTGTTGGGACTTGTCTTGTGACTTATGTGGCAATTTTGGAGAAGTCC	187011
Qy	507	ATAGGCTGTGCAAGAAGGACAGTCTTTGTGTTTGGTGAAA---TAGTCTGTAAATA	563
Db	187012	ATCAGGTGCTGAGAAGGATGTTCTGTTCTATTTTGGGTGAATGTTCTATAGATGA	187071
Qy	564	TCTCTAGTCCACTTGGTTTATGACATCAGTTACTCCAGATTTCTCTGTTTCGTTTT	623
Db	187072	ATGTTAGGTTCAATTTGAATCATACCTCTGTATGTTCTGTATTCTCTGTTTAG-TTTC	187130
Qy	624	TGTTGAGTAGCTTAACTTGTTGGAGAGAACTGGGTATTGAAGTAGCCACTATCTGTG	683
Db	187131	TGTCITGATGACATGCTCCAAAGGTAAGGTTGGGTGTTGAGTCTCCACTACTAATGTG	187190
Qy	684	T-GAGTCAATATGATTTTAGTGTAGCTGTGCTGTTGTTTTATGAACTTGGGTGACATT	742
Db	187191	TGCTGTTCATGTGTAATTAAGCTTTAGTCATGTTCTCTTTTACAAATGTGGTGTTCTTT	187250
Qy	743	GTCTTGTGTCATAGACATTAAGAATTTGCAATGCTCTTGTGGTGA-TTTTCTCTTTGATG	801
Db	187251	GCATTTGGGCGATAGATGTTCAGAAATGGGATGACCTCTTAGTGGATTTTTCTTTGATA	187310
Qy	802	CCTATGTAGTATTTCTTCCCAATCTCATCTGCTTAGTTTTCGGTTTAAGTCTATTAG-TCA	860
Db	187311	AGTATGAAGTGTACTCCCACTCTCTTTTGATTCCTTTTGGTGAAGTCTATTATTATA	187370
Qy	861	GATATTAATGACGTATCGGCTGCTCTTAGGGCCATTTCCTTAGAATATCTTTTCC	920
Db	187371	GTATTAGAATGGCTTAATCCAACTGTGTTCTGGGCTGTTGTTTGA AAACTTTTCC	187430
Qy	921	ATCCCTTTTACT---CTAAGGTGATGCTATCCATGGTAGGTTGCTTTTTTGGATGCAGCA	978
Db	187431	AGCCTTATGTTGAGACAGTGACTATTTTGTCTGATGATGTTCTTGATAGAGCA	187490
Qy	979	GTAGATGGAATCTGTTTTCATATCCATCTGTTTACCAGTATCTTTTTCTTAGAGAAATT	1038
Db	187491	GAATGATGAATCTGTTTATGCATCCATCTGTTTAGCCTTGTCTTTTAGTTGGGAATT	187550
Qy	1039	AAGATCATTGAGTCAATGATGTGAGAAATATCAATCAGCAGTGTGTTGGATCTTGTGT	1098
Db	187551	GAG-----TCTATTGAATTAAGATGTTAATGACCAATGATGTTACTTCTGT-187601	
Qy	1099	ATCTTGCACTGTGAAGT	1158
Db	187602	-----CTTTTGTGATATTGGTGGTGTGTTTGTGTGTGTGTGTGTGTGTGTGTGTGT	187655
Qy	1159	GT	1218
Db	187656	GT	187707
Qy	1219	TTGATTTTGGCTCGGAATTTATTTATTCATATTTTCTTGAATGTGGGTAAACATCTTT	1278
Db	187708	-----GAATATTTATTTCTGTGTGTTTCTTGTAGTGTAGTACCTTCATT-187752	
Qy	1279	AGATTGAAGTTTTTCTCTAGCCT---CTTTTAGGCTGCAATTTGAAGATAGATATCTTT	1336
Db	187753	GGGTGGAGTTTTCTTTTAGTATCTCTGAGGGCTGGGTATGTGGGAAGGTATTTGTTT	187812
Qy	1337	ACATCTGATTTTATCTTGAATGTCTTCTTTCTCCAACTATTGTGACAGAAGTTTTTC	1396

[illegible]

RESULT 29
AC101656/
LOCUS
DEFINITION

ACCESSION

AC101656

VERSION

AC101656.2 GI:30018174

KEYWORDS

HTG; HTGS PHASE1; HTGS DRAFT.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS

Biren, B., Nusbaum, C. and Lander, E.

JOURNAL

Mus musculus, clone RP23-306M24

REFERENCE

Unpublished

AUTHORS

2 (bases 1 to 267360)

AUTHORS

Biren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hages, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kellis, C., LaRocque, K., Lanazares, R., Landers, T., Lehotzky, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbak, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

Direct Submission

JOURNAL

Submitted (23-NOV-2001)

REFERENCE

Research, 320 Charles Street, Cambridge, MA 02141, USA

AUTHORS

3 (bases 1 to 267360)

Biren, B., Nusbaum, C., Lander, E., Abouelkell, A., Allen, N., Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hages, B., Galagan, J., Gardyna, S., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kellis, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupbak, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

Direct Submission

JOURNAL

Submitted (23-APR-2003)

COMMENT

Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Apr 17, 2003 this sequence version replaced gi:17060431.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L17152
 Center clone name: 306_M_24

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 47 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence.
 * as soon as it is available and the accession number will
 * be preserved.

1 606: contig of 606 bp in length
 * 607 706: gap of 100 bp
 * 707 1376: contig of 670 bp in length
 * 1377 1476: gap of 100 bp
 * 1477 2112: contig of 636 bp in length
 * 2113 2212: gap of 100 bp
 * 2213 2934: contig of 722 bp in length
 * 2935 3034: gap of 100 bp
 * 3035 3825: contig of 791 bp in length
 * 3826 3925: gap of 100 bp
 * 3926 4857: contig of 932 bp in length
 * 4858 4857: gap of 100 bp
 * 4859 5869: contig of 912 bp in length
 * 4958 5969: gap of 100 bp
 * 5870 6546: contig of 577 bp in length
 * 5970 6646: gap of 100 bp
 * 6547 7092: contig of 446 bp in length
 * 6647 7192: gap of 100 bp
 * 7093 8443: contig of 951 bp in length
 * 7193 8443: gap of 100 bp
 * 8444 8856: contig of 613 bp in length
 * 8244 8856: gap of 100 bp
 * 8857 10016: contig of 1060 bp in length
 * 8957 10016: gap of 100 bp
 * 10017 10736: contig of 620 bp in length
 * 10117 10736: gap of 100 bp
 * 10737 11306: contig of 470 bp in length
 * 10837 11406: gap of 100 bp
 * 11307 12849: contig of 1443 bp in length
 * 11407 12949: gap of 100 bp
 * 12850 13716: contig of 767 bp in length
 * 12950 13816: gap of 100 bp
 * 13717 14302: contig of 486 bp in length
 * 13817 14402: gap of 100 bp
 * 14303 15778: contig of 1376 bp in length
 * 14403 15778: gap of 100 bp
 * 15779 16884: contig of 1006 bp in length
 * 15879 16884: gap of 100 bp
 * 16885 18208: contig of 1224 bp in length
 * 16985 18208: gap of 100 bp
 * 18209 19291: contig of 983 bp in length
 * 18309 19391: gap of 100 bp
 * 19292 20316: contig of 825 bp in length
 * 19392 20316: gap of 100 bp
 * 20317 21379: contig of 963 bp in length
 * 20317 21379: gap of 100 bp
 * 21280 22611: contig of 1232 bp in length
 * 21380 22611: gap of 100 bp
 * 22612 23936: contig of 1125 bp in length
 * 22612 23936: gap of 100 bp
 * 23837 24576: contig of 640 bp in length
 * 23937 24576: gap of 100 bp
 * 24577 25551: contig of 775 bp in length
 * 24577 25551: gap of 100 bp
 * 25452 26907: contig of 1356 bp in length
 * 25552 26907: gap of 100 bp
 * 26908 28014: contig of 1007 bp in length
 * 27008 28014: gap of 100 bp
 * 28015 29141: contig of 1027 bp in length
 * 28115 29141: gap of 100 bp
 * 29142 30593: contig of 1352 bp in length
 * 29242 30593: gap of 100 bp
 * 30594 31784: contig of 1031 bp in length
 * 30594 31784: gap of 100 bp
 * 31785 31784: gap of 100 bp

* 31885	33263: contig of 1379 bp in length	Db	125647	ATGAGTGTCTGAGAAGAGGTATGTCGTTGTTATTTGGGTGAAATGTTCTATAGATGTA	125588
* 33264	33363: gap of 100 bp				
* 33364	34467: contig of 1104 bp in length	QY	564	TCTTAGGTCCACTTGGTTTATGACATCAGTTAGTCCAGCATTTCTCTGTTTGGTTTTT	623
* 34467	34567: gap of 100 bp				
* 34568	35942: contig of 1375 bp in length	Db	125587	ATGTTAGGTTTCAATTTGAATCAATCACTCTGTTTGGTTTCTGTTTCTGTTAG-TTTC	125529
* 35943	36042: gap of 100 bp				
* 35943	37569: contig of 1527 bp in length	QY	624	TGTTGAGATGACCTAACTGTTGGAGAGAAATGGGGTATTGAAGTAGCCACATCTGTTG	683
* 37570	37669: gap of 100 bp				
* 37670	39143: contig of 1474 bp in length	Db	125528	TGCTTTGATGACATGCCAATGTTAAGAGTGGGGTGTGAGGCTCCCACTACTAATGTG	125469
* 39144	39243: gap of 100 bp				
* 39244	41309: contig of 2066 bp in length	QY	684	T-CAGGTCAATATGTGATTTTACGTGTAGCTGTCTGTTTATGAACTTGGGTGACAT	742
* 41310	41409: gap of 100 bp				
* 41410	42879: contig of 1470 bp in length	Db	125468	TGGTGTTCATGCTGTGATTTAAGCTTTAGTCATGTTCTTTTACAAATGGGTGTTCT	125409
* 42880	42979: gap of 100 bp				
* 42980	44743: contig of 1764 bp in length	QY	743	GTGTTGGTGCATAGACATTAAGAAATTCGAATGCTCTCTGGTGA-TTTTCTTTGATG	801
* 44744	44843: gap of 100 bp				
* 44844	46754: contig of 1911 bp in length	Db	125408	GCAATTTGGGGCATAGATGTTTCAAGAAATGGGATGACCTCTTAGTGGATTTTCTCT	125349
* 46755	46854: gap of 100 bp				
* 46855	48072: contig of 1218 bp in length	QY	802	CCTATGTAGTATCTTCCCAATCTCATCTGCTTACTGTTTGGGTTTAAGTCTATTAG-TCA	860
* 48073	48172: gap of 100 bp				
* 48173	50344: contig of 2172 bp in length	Db	125348	AGTATGAAGTGACTTCCCAATCTCTTTGATGCTTTTGGTGAAGAGTCTATTATTATTA	125289
* 50345	50444: gap of 100 bp				
* 50445	53613: contig of 3169 bp in length	QY	861	GATATTAATAAGTGTATCGGCTTGTCTTTPAGGGCAATTTGCTTGAATATCTTTTCC	920
* 53614	53713: gap of 100 bp				
* 53714	94770: contig of 41057 bp in length	Db	125288	GTTATTAGAAATGGCTAATCCAACTTGTCTTGGGTCTGTTTGTGTTGAAAACCTTTTCC	125229
* 94771	94870: gap of 100 bp				
* 94871	178917: contig of 82047 bp in length	QY	921	ATCCCTTTTACT--CTAAGGTGATGCTATCCATGCTGCTTTTGGATGAGTCTTTTGGATGACGA	978
* 178918	179017: gap of 100 bp				
* 179018	267360: contig of 90343 bp in length.	Db	125228	AGCCCTTTATGTTGAGACAGTGAATTTTGTGTTGAGATGTTTCTTGTATAGACA	125169
FEATURES					
Location/Qualifiers					
1..267360	/organism="Mus musculus"	QY	979	GTAGATGGATCTGTTTTCATATCCATTCGTTCATCCAGTATCTTTTCTAGAGAAAT	1038
1..267360	/no_type="genomic DNA"				
1..267360	/db_xref="taxon:10090"	Db	125168	GAATGATGATTCCTGTTTATGCAATCCATTCGTAGGCTTGTCTTTTGTAGTTGGAAAT	125109
1..267360	/clone="RP23-306M24"				
1..606	/clone_lib="RPC1-23 Female Mouse BAC"	QY	1039	AGATCAATGAGTCATGATGTTGAGAAATATCAATGACGAGTGTGTTGGAATCTTTGT	1098
1..606	/note="assembly_fragment"	Db	125108	GAG-----TCTATTGATGTTAAGAGATGTTAATGACCAATGATTTACTTCTCTGT-	125058
707..1376	/note="assembly_fragment"	QY	1099	AUCTTGCACTTCTGAAAGT	1158
1477..2112	/note="assembly_fragment"	Db	125057	-----CTTTTGTGATATTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	125004
2213..2934	/note="assembly_fragment"	QY	1159	GT	1218
3035..3825	/note="assembly_fragment"	Db	125003	GT	124952
3926..4857	/note="assembly_fragment"	QY	1219	TTGATTTTGGCCTGGAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	1278
4958..5869	/note="assembly_fragment"	Db	124951	-----GAATTTATTTTCTGT	124907
5970..6546	/note="assembly_fragment"	QY	1279	AGATTGAAGTTTTTCTCTAGCCT--TCTTTAGGCTGTGCATTTGAGATAGATATCTTT	1336
6547..7360	/note="assembly_fragment"	Db	124906	GGTTGGAGTTTTTCTCTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	124847
7361..8080	/note="assembly_fragment"	QY	1337	ACATCTGATTTTATCTAGAAATCTTCTTCTCCAACTATTTGACAGAGAAAGTTTTTC	1396
8081..8800	/note="assembly_fragment"	Db	124846	AAATTTGGTTTTGTTGTGGAATATCTTGATTTCTCCACTATGGTGATCGAAATTTGC	124787
8801..9520	/note="assembly_fragment"	QY	1397	TAAGTGCAGTGTCTGGCCTGCATCTGTAGTCTCTCTGGAGTCTGTAGCATCTGTGCA	1456
9521..10240	/note="assembly_fragment"	Db	124786	TGGGTGTAGTGTCTGTGCTGACATCTGTGTTCTCTAGGGTTTCAAGACCTCTGCCA	124727
10241..10960	/note="assembly_fragment"	QY	1457	GGCCCTCTTACATTTGAGTTTCTATTGGAAAAGTCAAGTGTATTTCTATACATCTGC	1516
10961..11680	/note="assembly_fragment"	Db	124726	GGCCCTCTTCTGTTTTTAGAGTCTC-ATGTGATAAATGGGTATAAATTTGATAGTCTGC	124668
11681..12400	/note="assembly_fragment"	QY	1517	CTTTATGCTTAATTTGGTCTTTTTTCCCTTGCATCTTTTAAATTTCTTTCTTTCTAT	1576
12401..13120	/note="assembly_fragment"	Db	124667	CTTTATGCTTAATTTGGTCTTTTTTCCCTTGCATCTTTTAAAGTCTTTCTTTCTAT	124609
13121..13840	/note="assembly_fragment"	QY	1577	ACTTTTGTGATTTGATTTATTAATGACCTGTGGGAGTTTCTTTCCGTTCCAACTATTT	1636

Query Match	30.9%; Score 704.103; DB 1; Length 267360;
Best Local Similarity	65.1%; Pred. No. 0.18; Gaps 23;
Matches 1355; Conservative	0; Mismatches 574; Indels 153; Gaps
QY 270	TTCCCTCTCTCTCCAAACACTCTTATTTCTTGATTTCTATCTTGTGCTCAATTTTAACTC 329
Db 125880	TTCACTGAATTTTGAAGAGTCTTAAATTTCTTTTATTTCTTCCCTAA-----TCATTG 125825
QY 330	AGTAGTGAATTTTGGTTCCTCAATAGTTTCTGATTTCTGTTTCTGTTGTTCTGTTGTTG 389
Db 125824	AGTAGAGCCATTCAGTTTCTCAATATATGAGCTTTCTGTTTCTGTTTCTGTTTCTGTTG 125765
QY 390	TGTTATCTAGATTTAAAGTGTGTGGTGCAGATAGGA-CATAGAGATTTATTTCAATGT 448
Db 125764	AAG---TCCAGCTTTATTCATGTTGATCTGATAGATCCATGGGTTTATTTCAATCTT 125708
QY 449	CTTTATCTGCGAGACTGCTGTTTGTGTTTGAATATGATTTCAATTTGGAGA--GTTTC 506
Db 125707	CTTTATCTGTTGGACTTGCTTTGTGACTGATTTGGGCAATTTTGGAGAGGTTCC 125648
QY 507	ATAGGCTCTGACAAAGAGGTACAGTCTTTGTTGTTTGGTGA---TAGTCTGTAATA 563

Db	124608	ATAATTAGTGTTTTCGATTATTAATGATGGTGGAGATTTTCCTTTTGGTCCAAATCTGTTT	124549
Qy	1637	GGTGTTCCTATGCTCTCTGTACCTTGATAGGCAATCTCTTCTCAAGGTTAGGAAATTTT	1696
Db	124548	GGTGTCTGCTTACCTCTCTGTACCTTTATAGGCATCTCTTCTTTAGGTTAGGTAATTT	124489
Qy	1697	TCCTTTTTCG-GTTTCTCTGAAATATTTCCCTGCTTTTGACCTGC-CCTCTTCCCTTC	1754
Db	124488	TCCTCTATGATTTTGTGAGAGATGTTTCTCGGCCCTTTGAGCTGCAAAATATCTCCTTC	124429
Qy	1755	CTCTATTC-C-----TTTGGTTTTTGATAGTGTCTCTGGCTTCCCTGGATGTT	1801
Db	124428	TTCTATTCTCTGTTATCTTTAGGTTTGGTCTTTTCACACTGTGCCAAATTTTCTGGATGT	124369
Qy	1802	TTATGCGCTGATTATTTTAGACTTAACATTTTCTTTGACCAGGATATCCATTTCTTCTAT	1861
Db	124368	TTATGTCAATGTAATTTTCAGATTTTGCATTTTCTTGACTGACATATAATTTCTTCTAT	124309
Qy	1862	CTTGCTTCCAGTCGCTGAGATCTCTCTCTATCTCTTTGATTCTGTCACTGAGGCTTCT	1921
Db	124308	TGTAATCTTCTATCTTGAGATC-----TCTATCGCTGATCTCTGTTGATACTTAT	124254
Qy	1922	CTCTGAGGTTCCCTGTGGTCTTAATTTTTCATTTTCCAGATTTCTCTCAGTTTGGT	1981
Db	124253	GCCCTTATCTCTGCTCTCTCTCCTAGGTTTATATCTCCAGATGTGCTCAGGTTGTT	124194
Qy	1982	TTGTTTATTAATCTTATTTTCCACTTTCAGGTCCTGAAATGTTTTTACTCATTTTCCFCCA	2041
Db	124193	TTCTTTATTGCTTCTATTTCCCTTTTCAGGTGTGGACAGTTTTTATCATATCCTCACC	124134
Qy	2042	GTATT-----GTATT-----	2046
Db	124133	TGTTGATGTATTTTCTGTAATCTTATACTTATGTTGCTTACTTTTAAATGTCCTACC	124074
Qy	2047	-----TACATTTTCATAGGTTCTCTTAAATGATTTATCTATTCCTCTCAAGGACC	2098
Db	124073	TGTTTGAATGATTTTTCCTGTATTTCTTTAGAGATTTACTCATGTCCTCTTAAAGGCT	124014
Qy	2099	TTTTATGAATTCATAAAATGATGTTAAGTTCCTTGCTTGTGCTTCAGCTATGTTGCAAT	2158
Db	124013	TCGTCACTCTTATATAAATTACATTTAAGATCTCTCTTCTGTACTTGTTTTCAT-TAGGGT	123955
Qy	2159	TCTCAGGGCTATTCGTAATA-----GGGTTTTAGCAGGACATATGCTCTGCTGTTCT	2210
Db	123954	ATCCAGGGTTTGTGTAGGAGGACAGCTGGACTCTGTGGTATCATAGTCTCTCTGGTCT	123895
Qy	2211	TATGTCTGTGTTTTTGTGTTTGGCAATATAGACGCTCAGTTT	2252
Db	123894	TATTGATGTGTCTGAAACTGTCTTTTAGGAACTGTGTTT	123853

Db	117069	TCCTTATTTTATTTTACCACAGTGGTCATTGAGTAAGGGGGTATTCAGTTTCAATGAGT	117128
Qy	358	TTGTAAAGTTTTCTGTTGTTCTGTGTGTGTTGTTGTTATCTAGATATTTAAAGTGTGGTGGT	417
Db	117129	TTGCTAGAAATTCCTGTGTGTTCTGATATTAAGGT-----CAGCTTTAATCCATGTTGGT	117182
Qy	418	CAGATAGGACATAGAGTATTATTTCAAATATGCTTTTATCTCTCAGACCTTGCTTTGTTTT	477
Db	117183	CAGATAAGATACATGGGGTTATTTCATTTTATTATATGTGTGAGTTTTCGTTTCATGAC	117242
Qy	478	GAATAATGTATTCAAATTTTGGAGA--GTTTCATAGGGTGCTGACAAGAAAGTACAGTCTC-	533
Db	117243	TAAGTATATGATCAATGCTTCGAGAAGGTTTTTATGAGTGATGAGAAGGATATTCT	117302
Qy	534	TTTGTGTTTTGGTGAATAGTCTGTAATAATCTCTAGTCCACTTGTTTATGATCATCAG	593
Db	117303	TTTGTGTTTTGGGTGAATGTTCTGTGATGTCGTTAGATCCATTGTTTCAATACATCTA	117362
Qy	594	TTAGCTCCACATTTCTCTGTTGTTTTTTGTTTGAGATGACCTAACTGTGTGGAGAGAT	653
Db	117363	TTAGTTTCATTATTTCTGTGTTACTTT-----CTGCCCTGGATGAATCAG	117408

654 GGGTAATTGAAGTACCCACATCTCTGTGTGAGG-TCAATATGTGATTTAGCTGTAGC 712
Db |||||
117409 TAGGTGTTGAATCCGTACATTAATCTATGATGCGGTTCGATGTGTGATTTAAGCTCTAGT 117468
Qy |||||
713 TGTGCTGTTTATGAACCTCGGTGACATGTTGTTGGTGATAGACATTAAGAAATGCA 772
Db |||||
117469 CGTGTTCCTTTATGAATGTTGGTGCCTTACATTTGAGGTGTAGATACACAGATTAAG 117528
Db |||||
773 ATGTCCTCTGGTGGATTTTCTTTGATGCTATGTAGTATCTTCCCAATCTCATCTGC 832
Qy |||||
117529 ACTTCATCTGCTAGTTTTCCTT-TGAGTATGAAGTGTCTCCCACTCTCTTTAA 117586
Db |||||
833 TTAGTTTGGGTAAAGTCTA-TTAGTCAGATATTAATAAGTGTATCGGCTGTCTCT 891
Qy |||||
117587 TTAATTTGATCGAAGTCTATTTTAAATGATATAGATGGCTACACCAAGTTGCTTCT 117646
Db |||||
892 TAGGGCCATTTGCTTAGAATATCTTTTCCATCTTTTACTCTAAGGTGATGTCTATCCAT 951
Qy |||||
117647 TGGGTCCATTTTCTTGAATAATAATTTTCCAGCCTTTTACACTGAGGTAGTGCCTACCTTT 117706
Db |||||
952 GGT-----AGGTGTCTTTTGGATGACAGTAGGATGGATCTGTTTTCATATCCAT 1007
Qy |||||
117707 GTTGTGAGGTGTATTTTATATGACAGCAATGATGATCTGTTT-----AACA 117759
Db |||||
1008 CTGTTACCCAGTATCTTTTCTAGAGAAATTAAGATCATTCAGTCAATGTATGTGAGAAT 1067
Db |||||
1068 TATCAATGAGCAGTGTGTTGATCTTTGATCTTTGATCTGCACTTGTGAAGTGTGTGTG 1127
Qy |||||
117812 TATTAATGACCAATGATGTTTATTTCTGTTTATCTTGAATGTTGTTAGTAGTGTG 117871
Db |||||
1128 TGT 1187
Qy |||||
117872 TGT 117905
Db |||||
1188 TGTCTGT 1247
Qy |||||
117906 -----TAGGTACTTCCCTTCTGTTGTTGCTGCTGTTAGTATTTTCTT 117952
Db |||||
1248 TC-ATATTTCTGTAATGCTGGTACATCTTTAGATGAAGTCTTTCTCCTAG-CCCTC 1304
Qy |||||
117953 CTTTGTCTCTTGAGTGTAGTATCTTACCTCTGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 118012
Db |||||
1305 TTTAGGTCTGCAATTTGAAGATAGATATCTTTTACATCTGATTTTATCTTAGAATGCTTT 1364
Qy |||||
118013 TGCAGATTTGATTTGTGATACATATTTGCTTAAATTTGTTCTGTCAATGATATCTGT 118072
Db |||||
1365 CTTTCTCAACTATTTGTGACAGAAAGTTTCTAAGTGCAGTACTCTGGCTGACATCTG 1424
Qy |||||
118073 GTTCTCCATCTATGGTGATTTGACCAATTTTCTGGGTATAATGCTGTGAGCTGTCACTG 118132
Db |||||
1425 TAGTCTCTTGGAGTCTGTAGACATCTGTGAGGCGCTTCTTACATTTTGAATTTCTATT 1484
Qy |||||
118133 TGTCTCTCTTAGGCTGTGACACATCTGCCAGGCGCTTCTGATTTTAAATCTCTGTT 118192
Db |||||
1485 GAAAGTCAAGTGTAAATCTTAATCAATCTGCTTTTATGTTTAAATGCTTTTCTTCC 1544
Qy |||||
118193 -GAAAGTCAAGTGTAAATCTGATAGACCTGCTTTATATTTACTTGG-CCCTTTCC 118250
Db |||||
1545 TTGCATCTTTTAAATCTTTCTTTGTTCTATCTTTTGTGATTTTGAATTTATGCACT 1604
Qy |||||
118251 TTATAGCTTTTCAACATCTTTCTTTGTTCTGTGCAATTTAGTGTGTTTGAATTTGTGAT 118310
Db |||||
1605 GTGGGAGTTCTTTTCCGTCCTCAATCTATTGTGTTTGTATGCTTTGTACCTTGA 1664
Qy |||||
118311 GGGAGATTTTCTTTCTGTTGCA-CTTATTTGTTTCTGTAAGCTTCTGTATGTTA 118369
Db |||||
1665 TAGGATCTTTCTCAGGTTAGGAATTTTCTTTTGTGTTTCTTCTGAAATATTT 1724
Qy |||||
118370 TAGGTATCTTTCTTTTAGGTAGGGGAATTTTCTTCTGTGATTTTGTGAGATAATTT 118429
Db |||||
1725 CCTGCTTTTACCTGCTCTTCTTCCCTTCTCT-----ATTCCTTTGG 1768
Qy |||||

118430 ATGGACCATGTAGTTGGGAATCTTCTTTCTATTCCCATTTATCTAGTTAGTTT 118489
Db |||||
1769 TTTTGTGATAGTGTCTGTGGCTTCTGGATGTTTATGCTGGATATTTTAGACTTAAC 1828
Qy |||||
118490 TTTAGGACATATCCAGATTTCTGGATGTTTGTATCAGGAAGTTTGTAGATTAAAC 118549
Db |||||
1829 ATTTTCTTTGACCAAGGTATCCATTTCTTATCTTGTCTTCACTGCTGAGATTTCTCTC 1888
Qy |||||
118550 ATTTTCTTTGACAGATGATCAATTTCTTCTGTATCTTACACATGAGGCTCTCTC 118609
Db |||||
1889 TTCTATCTCTTGTATCTGTCTGAGCTGTCTCTGAGGTTCTGTGGTT-CTTAA 1947
Qy |||||
118610 TTTTATCTCTGTATCTGTGTTGGTGTGATGCTTCTATCTGTGTCTCTTCTTACCAAG 118669
Db |||||
1948 TTTTCTTATTTCCAGATTTCTTCTGATTTGGGTTTGTGTTTATTTATTTCTTATTTCCACTTT 2007
Qy |||||
118670 GTTTTCTATCTCCAGATTTCTTCAATTTGTGTTTCAATTTCTTCTTCTTCTTCTTCTTCT 118729
Db |||||
2008 CAGGCTCTGAAATGTTTACTCAT---TTTCTCCAGTATTTTACATTTTCTATAGTTTC 2064
Qy |||||
118730 TATGCTTGAACAGTTTACTCATTTTCTTCACTGTTTGTATTTGCTGCTGTATTTTC 118789
Db |||||
2065 TTTATGATTTTATTTCTTCTTCAAGACCTTTTATGATTTCAATAATGATGTT 2124
Qy |||||
118790 TTTAAGGATTTTATTTCAATTTCTTTTAAAGGCTCTATCATTTTCAAGTTTGTAT 118849
Db |||||
2125 AAGTCTCTGCTGTGCTTCACTGATGTT-CAATCTCAGGCGCTTATTTAATAGGTT 2183
Qy |||||
118850 AAGTCAATTTCTGCTTCACTGTTTGTAGGTTACTCAGGCTTGTGTAGCAAGA 118909
Db |||||
2184 TTAGCAGGACATTTTGTCTGTTTGTATGTC-TGTGTTTCTTGTGTCATATAGAC 2242
Qy |||||
118910 GCTAGGCTCTGTGTATGCTGCTGCTGATGTTGACTTATGTTCTTCTACTCTGCTTAGCC 118969
Db |||||
2243 GCTGAG-TTTGGGATGTTGTAATTTCTAGGTCGAT 2279
Qy |||||
118970 TTTCTGGTTTGTATGTTATTTAGGTTTAGATGCTGGT 119007
Db |||||

AC126499 211245 bp DNA linear HTG 20-NOV-2002
Rattus norvegicus clone CH230-202120, WORKING DRAFT SEQUENCE, 8
unordered pieces.
AC126499
AC126499.4 GI:25138312
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 211245)
REFERENCE
AUTHORS
Muzny, D., Metzker, M., Lee, A., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Altschuler, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Ayodeji, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankensburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Chen, A.,
Chacko, J., Chavez, D., Chen, G., Chen, Y., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escoto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebrgeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Quevara, W.,
Gunnarathne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, K., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hughes, M.,

Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, H., Jiang, S., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.B., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhwa, L., Loulseg, H., Lozada, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmood, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokelenah, O., Okwunigbo, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkoch, C., Popper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Pu, L., Puazo, M., Quiros, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, S., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G., and Gibbs, R.A.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished
2 (bases 1 to 211245)

Worley, K.C.

Direct Submission

Submitted (06-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 211245)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 20, 2002 this sequence version replaced gi:23269518. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GZDT
Center clone name: CH230-202L20
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 195571 bases at least Q40
Consensus quality: 195554 bases at least Q30
Consensus quality: 20612 bases at least Q20
Estimated insert size: 194363; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (See http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 75757: contig of 75757 bp in length
* 75758 75857: gap of unknown length
* 75858 80590: contig of 4733 bp in length
* 80591 80590: gap of unknown length
* 80691 105825: contig of 25835 bp in length
* 105826 106225: gap of unknown length
* 106226 131324: contig of 24699 bp in length
* 131325 131424: gap of unknown length
* 131425 202894: contig of 71470 bp in length
* 202895 202994: gap of unknown length
* 202995 204118: contig of 1124 bp in length
* 204119 204118: gap of unknown length
* 204219 205311: contig of 1093 bp in length
* 205312 205412: gap of unknown length
* 205412 211245: contig of 5834 bp in length.

FEATURES

source

1..211245
/organism="Rattus norvegicus"

/mol_type="Genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-202L20"

1..1069

/note="wgs contig"

1120..2375

/note="wgs contig"

3941..5860

/note="wgs contig"

197927..200300

/note="wgs contig"

200506..201503

/note="wgs contig"

201534..202894

/note="wgs contig"

Query Match 30.8%; Score 702.902; DB 1; Length 211245;
Best Local Similarity 67.9%; Pred. No. 0.15;
Matches 1362; Conservative 0; Mismatches 576; Indels 67; Gaps 28;

QY 292 TCTATTTCTTGAATTTCTCTGTTGTTTCTCTGTTGTTGTTTATCTAGATTAAGCTGT 411

Db 113733 TTATTTCTTTATTTCTGTTTGTGACCCATTTTAATTCAGTAGACAGTTGTTCAATTTCC 113674

QY 352 ATAAGTTTCTGAAGTTTCTGTTGTTTCTCTGTTGTTGTTTATCTAGATTAAGCTGT 411

Db 113673 ATGAGTTTATAGTTTCTATTTGTTCTGTTTATTTATGATAT---GCATTTAATCTGT 113618

QY 412 GGTGGTCAGATGAGCATAGATATTTATTTCAATTTGTCCTTTATCTGTCGACATGCTT 471

Db 113617 GGTGGTCTCATATAAATGCGAGGAATTTATTTCAATTTCTTTGTTGTTGACATGCTT 113558

QY 472 TGTTTTGAATATGATTTCAATTTTGGAGAGTTTTCATAGGGTCTGACAGAAGGTACAG 531

Db 113557 TGTACCAAGATGTAAT---TTTAGAAGAAGTTTCATGGGTGCGAGAGCAAACTAAAC 113500

QY 532 TC---TTTGTGTTTGGTGAATAGTCTGTAAATATCT---CTAGGTCCACTTGGTTATGA 587

Db 113499 ACTCTTTTGTGTTGGGTAAATGTTCTGTAATAATCTCTAGGTCCACTAGTTTAAAA 113440

QY 588 CATCAGTTAGTCCAGCATTTCTGTTTCGTTTCTGTTGTTGATGACCTAACTGTTGA 647

Db 113439 TGTCTGTAAGCTCCAGTGTGTTGTTCTGTTT-ACTTTTGTCTAGATGACCTGTCATTGT 113361

QY 648 GAGATGGGTATGGAAGTAGGCCACTATCTCTGTGTGA-GGTCAATATGTGATTTTAC 706

abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-15J3 is from the RP23-23 Mouse PAC Library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

FEATURES		Location/Qualifiers	
source		1. .22762	
		/organism="Mus musculus"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:10090"	
		/chromosome="X"	
		/clone="RP23-15J3"	
		/clone_lib="RP23-23"	
Query Match		30.7%; Score 699.902; DB 1; Length 227762;	
Best Local Similarity		69.1%; Pred No. 0.16; 511; Indels 59; Gaps 24;	
Matches 1276; Conservative 0; Mismatches 511; Indels 59; Gaps 24;			
Qy	294	TATTTCTTGATTTCTATCTTGGCTCAATTTTAACTCAGTAGTGAGTTGTTTGGTTCCAT	
Dd	168005	TATTTTCAATTTTCCCTGACTGAGCAATCATTTAGTAGAGAGTTGTTTCAGTTTCCAA	
Qy	354	AAGTTTGAAGTTTCTGTTGTTTCTGTTGTTGTTGTTGTTTCTAGATTAAGCTGTGG	
Dd	168065	GAGTTTGAAGCTTCTGTTGTTTCTGTTGTTGTTGTTGTTTCTAGATTAAGCTGTGG	
Qy	414	TGGTCAGATGACATAGAGTATTTTCAATTTCTTTTATCTGTCGAGACTTGTCTTTG	
Dd	168122	TAGTCTGATAGAGTACAGGGGTTGTTTCAATTTCTTCTGTTGAGACTTGTCTTT	
Qy	474	TTTTTGAATATGTAATCAATTTTGGAG--AGTTTCATAGGGTGTGCAAGAGGTACAG	
Dd	168182	TATTTGAGTATGCGTCAACTTTGAGAAAGTGTGATGAGTGTGAGAAAGGTCCAT	
Qy	532	TCT--TTGTGTTTTCGGTGAATPAGTCTGTAAATATCT--CTAGGTCCACTGGTTTATGACA	
Dd	168242	TATATTGCAATGGGTGTAATGTTTGTGAATATCTGTTATGTTTATTTAGTTTATTA	
Qy	590	TCAGTTAGCTCCAGATTTCTCTGTTTCTGTTTGGTTTGTGATGACCTAACTGTGGAGA	
Dd	168302	CCAATTAGCTCCAGATTTCTCTGTTTCTGTTTCT--AGACGATTTGTCATGGTGA	
Qy	650	GAATGGGGTATGAGTAGCCCACTATCTGTGTGTGA--GGTCAATATGTGATTTAGCTG	
Dd	168357	GAATGGGGTATTAACCTCCCTACTATCAATGTGTTACGATCAATGTGATTTAACTG	
Qy	709	TAGCTGTGCTTGTGTTTATGAACCTGGGTGACATTTGTTGGTGCATAGACATTAAGAAT	
Dd	168417	TGGTAGTTTTTGTGTTTACAGTGTGGGTGCGCTTGTGTTGGAGCATAGATTTAGAAT	
Qy	769	TGCATATGCTCTTGGTGTGA--TTTTTCTTTGATGCGCTATGATGATTTCTTCCCAATCTCA	
Dd	168477	TAAATATCATCTTTGATGATTTTCTCTGATGAATATGATGATCTATGTCCTCCCATCTC-	
Qy	828	TCTCTTAGTTTTCGGTTTAACTCTATTAGTACAGATTAATAAATGACTGTATCGGCTTGC	
Dd	168536	TTTCTCTAAATGUGGTTTGA--TATTTTGTAGATTAATAAATGGCTACACCGTCTTGC	
Qy	888	TTCTTAGGGCAATTTGCTTGAATAT--CTTTTCCATCTTTTACTCTAAGGTGATGCTA	
Dd	168594	TTCTAAGTCCATTTGGCATGAATATCTTTTCCCAACCAATTTACCTTTGAGTAAATGTTA	
Qy	947	TCC--ATGGTAGGTTGCTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATG	
Dd	168654	TACTTGTATGTTGAGTATGTTTCTCATATCAGTAAAGGATGGGTTCTGTTTCTGTACC	
Qy	1004	CATTTCTGTTACCCAGTATCTTTTCTAGAGAAATTAAGATCAATGAGTCAATGATGTTGA	
Dd	168714	CATTTCTGTTAATCTGTGTCITTTTATTTGGGAATGA-----GGCCATTAGTAATGA	

Qy	1064	GAATTATCAATGAGCAGTGTGTTGTGATTCITTTTATCTTGCACTTGGAAGTGTGTG	1123
Dd	168765	AAAGTATCAATGCAACAATGATTTAAATTTCTGATTTCTTACGATGTTGTGCAATTTG	168824
Qy	1124	TCGTG	1183
Dd	168825	TCGTG	168884
Qy	1184	TCGTG	1243
Dd	168885	AGAGAGAGAGG-----GAGAGAGTCTCTTACTTTGGTTTGTGTGTGTGTGTGTGTG	168938
Qy	1244	TTATTCATATTTTCTTGTAATGTGGGTAAACATCTTTAGATTGAAGTTTCTCTCTA--GCC	1301
Dd	168939	TTTCTG	168998
Qy	1302	TTCTTTAGTGTGCAATTTGAAGATAGATATTCTTTACATCTCATTTTATCTTTAGAAATGTC	1361
Dd	168999	TTCTGTAGACTGGATTTGTAGATGATATTCCTTTAAATTTGACTTTTTCGTGAATATA	169058
Qy	1362	TTCTTTCTCCAACTATTGTGTGACAGAAAGTTTCTTAAGTGCAGTAGTGTGCCCTGACAT	1421
Dd	169059	T---TTTCTCTATTGTGTGTGA-----TTTGTGTGAGTATTGTAGTCTGACATGGCAT	169108
Qy	1422	CTGTAGTCTCTTGGAGTGTGTAGACATCTGTGTGAGGGCTTCTTACATTTTGTAGTTCT	1481
Dd	169109	CTGTAGTCTCTTAGAATCTGTAGAAATCTGTGCCATGCCCTTCTAGCTTTTAGAGTCTCT	169168
Qy	1482	ATTGGAAGAGTCAGGTGTAATTTCTAATACATCTGCTTTTATATGTTAAATGTGCTTTT	1541
Dd	169169	GTT--GAGAAAGTCAGGCATAATTTCTAATAGTCTGTGTTTATATGTTACTTGTCTCT--TT	169226
Qy	1542	CCCTTGCATCTTTTAAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1601
Dd	169227	CCCTTGCAGGTTTAAACATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	169286
Qy	1602	ACTGTGGGAGTTTCTTTCCGGTCCAACTATTTTGGTGTGTTTGTATGCTTCTTCTTCTTCT	1661
Dd	169287	GATGGAGGA--TTTCTTCTGCTCACTATTTTGGTGTCTCTGTAAAGCTTCTTCTTCTTCT	169345
Qy	1662	TCATAGGCACTCTTTCTCAAGGTTAGGAAATTTTCTTTTGGTGTGTTTCTTCTTCTTCTTCT	1720
Dd	169346	TCATAGGCACT--TTCTTTAGGTTAGGAAATTTTCTTGTAGGATTTTGTGAAAAATA	169402
Qy	1721	TTTTCCCTCTTTTGTACCTGCC--TTCTTCCCTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCT	1777
Dd	169403	TTTTGAGGCTTTTGTAGCTGGGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	169462
Qy	1778	AGTGTCTCTGGCTTCTTGTGATGTTTATGCTCTGATTTATTTTAGACTTAAACATTTCTTCTTCT	1837
Dd	169463	GGCTTCCCATATTTCTCTGATATTTTGTGTGAGGAAATTTTACAATTAACA--TTTCTTCTTCT	169521
Qy	1838	GACCAAGGTATCCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1897
Dd	169522	GACCAAGGTATCCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	169581
Qy	1898	TTGTATTTCTGTCAGTGTGCTGCTGAGGTTCTTCTTGTGGTCTTCTTAAATTTTCTTCTTCTTCT	1957
Dd	169582	TTGTATTTCTGTTGGTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	169641
Qy	1958	TCAGATTTCTTCTGATTTGGGTTTGTGTTTATTAATTTCTTATTTTCCATTTTCCAGTCTCTGA	2017
Dd	169642	TCAGAAATTAACCTGCTTGTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	169701
Qy	2018	AATGTTTTTCTCATTTTCTCTCCAGTATTTTACATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	2077
Dd	169702	ACAGTTTATTTGTATGGGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	169761
Qy	2078	TTCTTTCTCTTCTCAAGGACCTTTTATGAAATTCATAAATATGATGT	2123
Dd	169762	TTGTTTTTCTTAGATTAATAAATAATTTCTCATCAATTAATTTATTTT	169807

QY 450 TTTTATCTGTGAGACATCTGTTGTTTGAATAATGTAATCAATTTGGAGAGTTTCATA 509
Db 206083 TTGTGCTGTTGAGCTTGATCT-----TTGGAAGAGTTTCATG 206044
QY 510 GGGTCTGACAGAAGAGTACA-GTCTTTGTGTTTGGTGAATAAGTCTGTAATCTCT 569
Db 206043 ATGTGCTAACAGAAGATATATTTTATGTTTCAGGTGAATGTTTGTGAGTATGT-T 205985
QY 569 AGGTCCACTGGTTTATCACATCAGTTAGTCTCCAGCATTTCTGTTTTCGTTTTCGTT- 627
Db 205984 AGGTCCATTTGGTTTACATATTTGTTA-CTACAGCATTTCTAGATTCAGTTTAGTTT 205926
QY 628 ---GAGATGACCTAACTGTTGAGAGAGATGGGTATTGAAGTAGCCACATCTCTGTGT 684
Db 205925 TTGTCTAATGATGCTATTGGCAATTTGTGGAATTCAGATCTCTCATTATCAGGGGTG 205866
QY 685 GA-GGTCAATATGATTTTACCTAGCTGCTGCTGTTTATGAATCTGGTGACATTTG 743
Db 205865 GAGGTCAATATG---TTTAAAGCTGTAGTGTGTTTCTTTTAAAGATCTGGTG-CCTTA 205810
QY 744 TGTGTTGTCATAGACATTAAGAAATGCAATGCTCTCTGTTGGATTTT---CCTTTGAT 800
Db 205809 TGCTTGGAGGATAGATGTTAAATAATGAATGCTCTCTGTTTCTTTTTCCTCCCTGAT 205750
QY 801 GCCTATGTAGTATCTTCCCAATCTCATCTGCTTAGTTTGGGTAAAGTCTA-TTAGTC 859
Db 205749 GAGTATGTAGTGTCTCTTTATCTCTCTCTGATCAATTTTGATTTGAAGTCTACTTTGTC 205690
QY 860 AGATATTAATAATGACTGTATCGGCTGCTCTTAGGGCAATTTGCTTAGAATATCTTTT- 918
Db 205689 AGAATTAATAATGACTGCACAGCTGCTCTTAGGTTCAATTTGCTTGCATATCTTTT 205630
QY 919 --CCATCTTTTACTTAAGTGATGCTATCC---ATGSPAGTGTGCTTTTGGATG 973
Db 205629 TCCACATTTTACCTCGAGGTGATGCTATCTCTAGATG-TAAAGTATATTTCTTAGATC 205570
QY 974 CAGCAGTAGGATGA-----TCTGTTTTCATATCCATCTCTGTACCCAG 1018
Db 205569 CAGCAGAAAGATGATCTTTTATGTTGTTTGTATTTTATTTCTGTTAAATCTG 205510
QY 1019 TATCTTTTCTAGAGAAATTAAGATCAATGAGTCAATGATGAGAGATTAATCAATGAGC 1078
Db 205509 CGTCTTTTATG-----GAGAAATGACCATTAATGTCGAGAGATTAATGATC 205459
QY 1079 AGTGTGAGTATCTGTTATCTGACATCTGCACTGTGAAGTGTGTGTGTGTGTGTGT 1138
Db 205458 AGTGTGTTGATTTGCTATTTGTTGTTGT-----TGTAGTAGTA 205415
QY 1139 GT 1198
Db 205414 TGGTGTGATGCTGGTGTGTGTGAACGTGGAGGTGGAGGTGTGTGTGTGTGTGTGT 205355
QY 1199 TCTGTGTTCTCTCCCTCTTTTGAATTTTGGCCTGGAATTAATTTATTTATTTCT 1258
Db 205354 TGTGTGTCTGTTTCCAACTTCACTTGTGCTGTATGGGTGTTTATTTCTTGTCTTTAT 205295
QY 1259 TGAATGTGGTAAACATCTTTAGATGAGTTTCTCTAGACCTCTTTTAGGTCTGCAAT 1318
Db 205294 TGGGTATGTTAAATGCTTTAGATTTGAGTTTCTCTAGCTCTGCTA---TAGAGTA 205238
QY 1319 TGAAGATAGATATTTCTTTACATCTGATTTATCTTAGAATGCTTTCTTCTCCAACTAT 1378
Db 205237 GATTCCTAGATATTTGCTTACATTTGGTTTATCATTTGATGATGCTTATTTCTGCTAT 205178
QY 1379 TGTACAGAAAGTTTTTCTAAGTGCAGTAGTCTGGCCTGACATCTGTAGTCTCTTGGAGT 1438
Db 205177 TGTGATTAAGGTTTTGCTGGGTATGATGATGCTAGGTGGCACCATATGATGCCTTAGAAT 205118
QY 1439 CTGTAGCACATCTGTGAGGGCCCTTCTACATTTTGTAGTTTCTATTGGAAGATGTCAGTG 1498
Db 205117 CGGACGACATCTGTCCAGGCCATCTGGCCTTTAGAGTCTCCATTT-GAGNAGTTGGAAG 205059

QY 1499 TAATTCTAATACATCGCCTTTATATGTTAAATGGTCTTTTTCCTTCGATCTTTTAAT 1558
Db 205058 AAATTTCTAAAATGCTGCTCTTGACAGGTGACTTGGTCTTTTGGCTTGCACTTTATA- 205000
QY 1559 ATTCTTTCTTTGTTCTATATCTTTTAGTATTTAGATTTATATGACATCTGCGGAGTTTCTTT 1618
Db 204999 --TTTTTCTTTGTTCTATATTTAATGTTTATGATTTATGTCGCAAGAAATTTCTT 204942
QY 1619 TTCGGGTCCAAATCTATTTCGGTGTGTTGATGCTTCTTGTACCTTTGATAGCATCTCTTTC 1678
Db 204941 TTCTGCTTAGTCTATTTTCTG---TATGTGCTTCTATATCTTTAGAGTCTTCTCTTC 204885
QY 1679 TCAAGGTTAGGAAATTTTCTTTTGGTTTTCTTGAATAATTTTCCCTGCTTTTGACC 1738
Db 204884 ATTAGGTTAGGAAATTTTCTTTTATCGTTTACTGTAATAATTTTCTGTGGCTTGACC 204825
QY 1739 TG-----CCTTCTTCCCCTTCTCTATTCTTCTTGGTTTTTTCATAGTCTTC 1785
Db 204824 TGTTCCTTCTCTTCTTCTTGTACCTATCATCTTTGATTTTATCTTTTATCTCTCTGTTCC 204765
QY 1786 TGGCTTCTCTGGATGTTTATGCTCGATTAATTTTAGACTTTAAACATTTTCTTTGACCAAG 1845
Db 204764 AGATTTCTCTGGATGTTTAAATGCCAGGAGT-TTTTAGATTTAACTTTTTTTTG-TCAGG 204707
QY 1846 TATCCATTTCTCTATCTTGTCTTCTACTGCTGAGATTTCTCTCTCTATCTCTGTATTC 1905
Db 204706 TATCCATTTCTCTATTTGTGCTCTCAATGCAATAGATTTCTCTTTTATCTCTCTGTTAT 204648
QY 1906 TGTCACTGAGGCTTGTCTGAGGTTCTCTG-TTGGGTTCTTAATTTTTCATTTCCAGAT 1964
Db 204647 TGTAGTGAAGCTTGTCTCTGAAGTTCTCTGTTAAGTTCCTATATTTTTCATTTTATGAT 204598
QY 1965 TTCCTTCAGTTGGGTTTTGTTTTTATTAATCTATTTCCACTTTTCAGTCTCTGAAATGTTT 2024
Db 204587 AGCTCTCAGTTGTGTTTTCTTTTATGATTTGATTTTTCAGGCTCTGAGCTGTCT 204528
QY 2025 TACTCATTTTCTCCAGATTTTACATTTTTCATAGTTTCTTAAATGGAATTTATTCATTT 2084
Db 204527 TATACA-TTTCCTTCTATTGTTGTGTTTTTACAGATTTTCTTTAAGAGACTTATAAATTT 204459
QY 2085 CCTTTTCAAGGACCTTTTATGAATTCATAAAATGATGTTAAAGGTCCTTTCCTTTGCTT 2144
Db 204468 CCTTTTAAAGGACCTCTATCACACTCATAAAGCTATTTTAAAGTCATTTCTATGCT 204409
QY 2145 CAGCTATGTTGATTT-----CTCAGGCGCTATTTAATAGGTTTTAGCAGGAGCA 2195
Db 204408 AAGCTATGTTTCTAGTATTCAGGACAGTTTGTGTAGAGTTGCTGGGCTCTAGTGGGACC 204349
QY 2196 TATTGCTCTGTTGTTTATGTTGTTGTTTCTTTGGCATATAGACGCTGAGTTTGGG 2255
Db 204348 TATTGCTCTGTTGTTGTTGTTGTTGTTTATGCTGATGCTAGCATCAGGTTTGAG 204289
QY 2256 ATGATTTGAATTTCTAGGTGCT 2276
Db 204288 AAGATTGCATTTCTAGGTGAT 204268

RESULT 35

AC115335

LOCUS

DEFINITION

AC115335

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

AC115335 243576 bp DNA linear HTG 13-NOV-2002
Rattus norvegicus clone CH230-202N2, *** SEQUENCING IN PROGRESS
***, 3 unordered pieces.

AC115335 4 GI:24941844
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
Rattus norvegicus (Norway rat)

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

1 (bases 1 to 243576)

Muzny, D.Marie, Metzker, M.Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,

Db	25246	TGCTTGGAGATAGATGTTTAAAAATTGAAATGCTCTTGGTTTTTTTTTCCCGCTTGAAT	25187
Qy	801	GCCTATGTAGTATCTTTCCCAATCTCATCTGCTTAGTTTGGGTTTAAAGTCTA-TTAGTC	859
Db	25186	GAGTATGTAGTCCCTTCTTAATCTCTCTGATCAATTTTGATTTGGAAGTCTACTTTGTC	25127
Qy	860	AGATATTAATAAGACTGTATCGGCTTCTCTTAGGGCAATTTGCTTAGAATATCTTTT- 918	
Db	25126	AGAAATTAAATGACTGCACCGACTTGCTTCTTAGGTTCAATTTGCTGCAATATCTTTT	25067
Qy	919	--CCATCCTTTTACTCTAAGGTGAUCTATCC---ATGGTAGGTGTCTTTTGGGATG 973	
Db	25066	TCCCAATTTTACCCCTGAGGTGATCTATCCTAGATGTTTAAAGTATATTCTTAGATC	25007
Qy	974	CAGAGTAGATGGA-----TCCTGTTTTCATATCCATTCCTGTATTACCAG 1018	
Db	25006	CAGCAGAAAGATGGATTCCTTTATCTTGTGTGTGATTTTATTTATTTCTGTAAATCTG	24947
Qy	1019	TATCTTTTCTAGAGAAATTAGATCATTTAGATGATTTGATGTGAGAAATTATCAATGAGC	1078
Db	24946	CGTCTTTTATG-----GAGAAATGTGACCATTTAATGTGGAAGAAATTAATGATC	24896
Qy	1079	AGTGTGTGGAATCTTGTTATCTGCATTTGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTG 1138	
Db	24895	AGTGTGTGTGATTGTGCTATTTTGTGTGT-----TGTAGTAGTGA 24852	
Qy	1139	GTG 1198	
Db	24851	TGCTG 24792	
Qy	1199	TCGTGTCTCTCCCTCTTTTGATTTTGGGCTGGAATTATTTATTTATTCATATTTTCT 1258	
Db	24791	TGTGTGTCTGTTTCCAACTTTCATCTGCTGGTATGGGTTGTTTATCTTGTCTTTTAT 24732	
Qy	1259	TGAATGTGGGTAACTCTTTAGATTTGAAGTTTTTCTCTAGCCCTCTTTTAGGCTGTCAAT 1318	
Db	24731	TGGGTATGTGTAATGTCTTTAGATTGGAGTTTTCTCTTAGCTCTTGCTA---TAGAGTA 24675	
Qy	1319	TGAAGATAGATATCTTTTACATCTGATTTTATCTTAGAATGTCTTTCTTCCCAACTAT 1378	
Db	24674	GATTCCTAGATATGCTTACATTTGTTTTTATCATTTGAATGTCTTATTTTCTGTCTAT 24615	
Qy	1379	TGTGACAGAAAGTTTTCTTAAGTCAGTAGTCTGGCTGCACATCTGTAGTCTCTTGGAGT 1438	
Db	24614	TGTGATTTGAAGGTTTTGCTGGGTATGATAGTCTAGGCTGGCACTATGATGCTTAGAAT 24555	
Qy	1439	CTGTAGCACATCTGTGCAGGCTCTCTACATTTTGTAGTTTCTATTGAAAGTCAAGTG 1498	
Db	24554	CGGCAGACATCTGTCCAGGCCATCTCGGCCTTTAGTCTCCATTT-GAGAAGTTGGAAG 24496	
Qy	1499	TAATCTCAATACATCTGCCCTTATATGTTAATTTGTTCTTTTTTCCCTTGCATCTTTTAA 1558	
Db	24495	AAATCTCAAAATGTCTGTCTTGACAGTGACTTGTCTTTTTTGCTTGCACCTTTTATA- 24437	
Qy	1559	ATTCCTTTCTTTGTTCTAATCTTTTAGTGATTTGATTTATATGCACTGTGGGGAGTTTCT 1618	
Db	24436	--TTTTTCTTTGTCTGTATTTAATGTTTAGATTTATGTGTCGCAAAAGGAATTTCTT 24379	
Qy	1619	TTCCGGTCCAACTATTTTGTGTGTTTGATGCTTCTGTACCTTGATAGGCATCTCTTTC 1678	
Db	24378	TTCTGTCTAGTCTATTTGCTG---TAGTGCTTCTATATCTTTAGAGTCTTCTCTTC 24322	
Qy	1679	TCAGGTTAGGAAATTTTCTTTTTTTTGGTTTTCTTGAAATAATTTTCCCTGCTTTTGACC 1738	
Db	24321	ATTAGGTTAGGGAATTTTCTTTTATCGTTTTACTGTAAATATTTCTGTGGCTTTGACC 24262	
Qy	1739	TG-----CCTCTTCCCTTCCCTCTATCTCTTTGGTTTTTGCATAGTGTCTC 1785	
Db	24261	TGTTTCTTCTTCTACTTTGTACCTATCAATCTTTTGATTTTATCTTTTCATAGTGTCCC 24202	
Qy	1786	TGGCTTCCCTGGATGTTTTATGCTCGGATTAATTTAGACTTAACATTTTCTTGACCAAGG 1845	

Db	24201	AGATTTCCTGGAGTGTAAAGCCAGAGT-TTTTAGATTAACTATTTTCTTTTGG-TCTAGG	24144
Qy	1846	TATCCATTTCTTCTATCTTGTCTTCACTGCCGAGATTCCTCTCTCTATCTCTCTGTTATTC	1905
Db	24143	TATCCATTTG-TCTATTGTGTCTCAATGCAATTAGATTCCTCTCTTTATCTCTGCTGTTT	24085
Qy	1906	TGTCAGTGAGGCTTGCTCTGAGGTTCCGTG-TTGGGTCTTAAATTTTTCATTTCCAGT	1964
Db	24084	TGTTAGTGAGGCTTGCTCTGAGGTTCCGTGTTAAAGTTCCTATATTTTTCATTTTATAG	24025
Qy	1965	TTCTCTTCAGTTTGGGTTTTGTTTAAATTTCTATTTTCCACTTTTCAGGTCCTGAAATGTTT	2024
Db	24024	AGCTTCAGTTTGTGTTTCTTTATTTGATTGTTATTTGATTTTCAGGTTCTGAGCTGCT	23965
Qy	2025	TACTCATTTTCTCCGAGTATTTACATTTTCAATAGTTTCTTTTAATGATTTTATTCATTT	2084
Db	23964	TATACA-TTTCTTTCTATTGGTTGTGTTTTCACAGATTTCTTTAAGAGACTTATAAATTT	23906
Qy	2085	CCCTCTCAAGGACCTTTTATGAATTCATAAAATGTATGTTAAGGTCCTTGCTCTGCTT	2144
Db	23905	CCCTTTTAAGGACCTCTATCACATCAAAAGGCTATTTTAAGGTCATTTATCCATGCT	23846
Qy	2145	CAGCTATGTTGCAT-----CTCAGGGCTATTGTAATGAGGTTTTCAGGAGACA	2195
Db	23845	AAGCTATGTTTCAGTATTCGAGCAGATTTGTGCTGAGTTGCTGGCTCTAGTGGGACC	23786
Qy	2196	TATTGTCCTGGTTGTTATTGTTCTGTGTTTCTGTTTGGCATATAGACGCTCAGTTTGGG	2255
Db	23785	TATTGTCCTGGTTGTTATGCTCATCTGTTGTTTATGCTGATGCTAGCATCAGGTTTGA	23726
Qy	2256	ATGATTGTAATCTAGGTGCT	2276
Db	23725	AAGATTGCACCTCTAGGTGAT	23705
RESULT 37			
AC125973			
LOCUS			
DEFINITION	AC125973	228302 bp DNA linear HTG 13-MAY-2003	
		Rattus norvegicus clone CH230-64N5, *** SEQUENCING IN PROGRESS ***	
		5 unordered pieces.	
ACCESSION	AC125973.3	GI:30579661	
VERSION	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.		
KEYWORDS	Rattus norvegicus (Norway rat)		
SOURCE	Rattus norvegicus		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
REFERENCE	1 (bases 1 to 228302)		
AUTHORS	Muzny,D,Marie, Metzker,M,Lee, Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,C., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Blawie,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Duan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,W., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M., Hollins,B., Howells,S., Kulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,S., Johnson,R., Jolivet,A., Karpthy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C., Labow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,		

Lorensheva, L., Loulsegged, H., Lozano, R.J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mahoney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K.,
 Nwokwelen, O., Okwono, G., Olarpunsgoon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
 Puato, M., Quirroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, X., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
 Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sibson, I., Sitter, C.D., Smales, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wiczyski, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, P., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstein, G. and Gibbs, R.A.

TITLE JOURNAL REFERENCE

Unpublished
2 (bases 1 to 228302)

Worley, K.C.

Direct Submission

Submitted (02-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 228302)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

On May 13, 2003 this sequence version replaced gi:22857433.

The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GLNF

Center clone name: CH230-64N5

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 210267 bases at least Q40

Consensus quality: 214022 bases at least Q30

Consensus quality: 217194 bases at least Q20

Estimated insert size: 221667; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 5 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 215461: contig of 215461 bp in length
 * 215462 215561: gap of unknown length
 * 215562 218013: contig of 2452 bp in length
 * 218014 218113: gap of unknown length
 * 218114 219376: contig of 1263 bp in length
 * 219377 219476: gap of unknown length
 * 219477 221346: contig of 1870 bp in length
 * 221347 221446: gap of unknown length
 * 221447 228302: contig of 6856 bp in length.

FEATURES source

1. 228302
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-64N5"

Query Match 30.6%; Score 697.602; DB 1; Length 228302;

Best Local Similarity 66.1%; Pred. No. 0.17;

Matches 1282; Conservative 0; Mismatches 554; Indels 104; Gaps 20;

QY 270 TTCCCTCTCTCTCCAAACACCTTCTATTTCTTGATTTCTATCTTGCTCATTTTAACTC 329

DB 84182 TTCCCTGAATCTAGAAAGTTTAAATTTCTTTATTTTCTTCCACCCAGAGATCATTG 84241

QY 330 AGTAGTGTGTTGGTTCCTTCCATTAAGTTTGAAGTTTCTGTGTTCTGTGTTGTTG 389

DB 84242 GATAAAGAGTTGTTCTATTTCCTCGAGTTTGAAGCTTTTGGTGTGTTGTTGAAG 84301

QY 390 TTGTTATCTAGATTTAAAGCTGTGGTGCAGATAGACATAGATATTTCAATTGTC 449

DB 84302 -----TCAGCTTAAATCCATGGTGGTCTGATGATACAGGCATTTTCAATTTC 84355

QY 450 TTTTATCTGTGAGACTTCTTGTGTTTGAATATGATTTCAATTTTGAGA--GTTTCA 507

DB 84356 TTCTATCTGTGAGGCTTGTGTTTGGCCCAAGTATATGTTTGGAGAACGTTTCA 84415

QY 508 TAGGGTCTGACAGAGCTGACAGCTTTGTTGTTTGGTGAATAGTCTGTAATATCTC 567

DB 84416 CAAGGTGCTGAGAAGAGTATTTCTTTGTTGTTG-----GGTGATATTG 84463

QY 568 TAGGTCCACTTGGTTTATGACATCAGTACCTCCAGCATTTCTCTTTTCGTTTTTGT 627

DB 84464 TTGGGTCCTTTGATTTCATAAATCAGTTAGTTCATTATTTCTGCTTAGTTTTGTCT 84523

QY 628 GAGATGACCTAATCTTTGGAGAGATGGGGTATTGAGTAGCCCATCTCTGTGTGAG 687

DB 84524 GGATAACCTGTTCAAGTTGTGAGAGTAGGGTGTGAAGTCTCTCACTATTAAATGTCGG 84593

QY 688 G-TCAATATGTGATTTTAGCTGTAGCTGTGCTTGTGTTTATGAACCTTGGTGACATTTGT 746

DB 84584 GTTTGACGTGTGATTTATCTTTTAGCAATGTTCTTTAACAATATGTTGGTGGCTTGCAT 84643

QY 747 TTGGTGATAGACATTAAGATTTGCAATGCTCTTTGGTGA--TTTCTTTTATGACCTA 805

DB 84644 TTATGGCCTAGATGTTTCAAGACTGAGATAAATCTCGGTATATTTTCTTTGATGACTA 84703

QY 806 TGTAGTATTCTTCCCAATCTCATCTCTTAGTTTGGGTTTAAAGTCTA-TTAGTCAGATA 864

DB 84704 TGAAGTGTCTTTGCTGCTCTTTTGTAGTTTGTGTTTAAAGTCTATTTTATTAGATA 84763

QY 865 TTAATAATGACTGATCGGCTTGTCTTCTTAGGGCCATTTCTCTTAGAATA-TCTTTTCCATC 923

DB 84764 TTAGCATGGTACTTTCAGCTTGTCTTCTTGGGACCATTTCTTCTTGGAAACCTTTTCTTACC 84823

QY 924 CTTTACTCTTAGGTGATGCTATC---CATGGTAGTTGTCTTTTCTGGATGCAGCAGT 980

DB 84824 CCTTACTCTAAGGCAATGCTATCTTTCTTCTGCTAGGTGCTGTTCTAGTATGCAGTAGA 84883

Qy	981	AGGATGAGACTTGTGTTTTCATATCCATCTCTGTACCAAGTATCTTTTTCTAGAGAAATTAA	1040
Db	84984	ATGCTGGGTTCTGTTTTTACAGCATATCTATAATCTGTCTCTTTTTATTTGGGGGATTGA	84943
Qy	1041	GATCATTGAGCTCAATGATGTTGAGATTATCAATGAGCAGTGTGTGTGGATTCCTGTTAT	1100
Db	84944	G-----TCCATTGATCTGAGAGATTAATGATGACAGTGAATGCTAAATTTCTGATAT	84995
Qy	1101	CTTGCACTTGTGAAGT	1160
Db	84996	TTTGATTTGGTGCTGGT	85055
Qy	1161	GT	1220
Db	85056	GTGTG-----CACGGCACGGGCTCTCCCTGTTTTTG	85088
Qy	1221	GAATTTTGGCTCGAAATATTTATTAATCATATTTTCTTTGAATGGGTAACATCTTTAG	1280
Db	85089	CTCTTGCTGTATGGAATCTTATTTCCCTGTGTTTTCTTTGGATGTGTTATATCTCCTTGG	85148
Qy	1281	ATTGAGATTTTTCTCCTAG--CCCTCTTTAGTCTGCAATTTGAAGATAGATATCTTTAC	1338
Db	85149	GTTGGAGTTTCTTTTGAGTATTTCTGTGAAGCTCGATTTGTGGATAGATATCTGTTAA	85208
Qy	1339	ATCTGATTTTATCTTAGAATGTCTTTCTTCCAACTATTGTGACAGAAAGTTTTTCTTA	1398
Db	85209	ATTGTGTTTTGTCTTGAAATTAATGTGTTTTCTCTATCTATGATGATTGAGAGTTTTGC--	85266
Qy	1399	AGTGCAGTAGCTGGCCCTGACATCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGCAGG	1458
Db	85267	--TCTAGTAGTCTATGTTGAGATCTGTGATGTTTTTAGAGGCTACAGACATCTGTTCAGG	85324
Qy	1459	GCCTTCTTACATTTTTCAGTTTCTATTGGAAAAGTCAGGTGTAATCTTAATACATCTGGCT	1518
Db	85325	CCCTTCTAGATTTTAGAGTCTCTGTGG--GAAGTTGGGTGTAAATCTGTATATCTGGCT	85393
Qy	1519	TTATAGTTAAATGGTCTTTTTTCCCTTGTGCATCTTTTAATAATCTTTCTTTCTCTATAC	1578
Db	85384	TTGTATGTT--AGTGGGCTCTTTCCCTGTGAAGCTTTTAATATTTTCTTCTCTCTGTAC	85442
Qy	1579	TTTTAGTGATTTGATATTAATGACGTGGGGAGTTTCTTTCCGGTCCAATCTAATTGG	1638
Db	85443	ATTTACTGTTTTGATGATTTGTGA--GCTGAGAGATTTCTTTCTGATCTAATCTAATTGG	85501
Qy	1639	TGTTTTGTATGCTCTTTG-----TACCTTGATAGGCATCTCTTCTCAAGGTTAGGAA	1691
Db	85502	TGTTCTGTAAGTCTTGATGTTTACTTTTACAGGCACTCTTCTTTAGGTTTGTAA	85561
Qy	1692	ATTTTTCTTTTTTGGTTTTCTTGAAAATATTTTCCCTGCTTTTGACCT--GCCTTCTTCCC	1750
Db	85562	TGTTTTCTTCTATGATCTGCTGAAAATATTTTCTAGTCTCTTGGAACTGGCACTCTTTAG	85621
Qy	1751	CTTCTCTATCTCTTGGTTTTTTCGATGTGTCCT-----GGC	1789
Db	85622	CTTCTCTATTCCTATTAATTTCTTAGCTTAGTCTCTTTATAGTATCCCTAGATGTATGTA	85681
Qy	1790	TTCTCGAATGTTTTATGCCCTGGATATTTTAGACTTAAATTTTCTTTGACCAAGTATC	1849
Db	85682	TCCTTGGAATTTTGTGTGACAGGAATCTTTTAGAATTTAAATTTTCTCTGACTGATGATC	85741
Qy	1850	CATTCTTCTATCTGTCTTCACTGGCTGAGATTCCTCTTATCTCTTTGTATTCTGTC	1909
Db	85742	AAATTTCTTCTATAGCATCTTCTAGACTGAGATTCCTCTTATCTCTCTGTGTTCTGTT	85801
Qy	1910	AGTGAGCTTTGTCTCTGAGGTTCTGTGTGGGTTCTTAATTTTTTCAATTTCCAGATTTCT	1969
Db	85802	GGTGATGCTTGCATCTGTGTTCCCGTCTCTCCCTAGGTCTTCACTCTCAGCATCTCC-	85860
Qy	1970	TCAGTTTGGTTTTGTTTTATTAATCTTAATTTCCACTTTCAGTCTCGAATCTTTTACTC	2029
Db	85861	-CAGTTGTGTTTTCTTTATGTCTTATTTTCAATTTTAGGTCCTTGAACAGCTTTATTT	85919
Qy	2030	ATTTTCTCTCCAGTATTACATTTTTCATAGGTTTCTTTAATGGA--TTTTATTCAATTTCT	2087

[illegible]

RESULT 38
AC134939

LOCUS AC134939 183078 bp DNA linear HTG 15-NOV-2002
 DEFINITION Rattus norvegicus clone CH320-375a21, WORKING DRAFT SEQUENCE, 3
 unnumbered pieces.
 ACCESSION AC134939
 VERSION AC134939.2 GI:25007201
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT;
 SOURCE Rattus norvegicus (Norway rat)

ORGANISM *Rattus norvegicus*
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

1 (basses 1 to 183078)

Muzny, D. Marie, Metzker, M. Lee, Abramzon, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebech, V., Ayadi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Bisswalo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, B., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Benson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhewa, L., Loulsegged, H., Lozado, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mathewey, S., McLeod, M. P., McNeill, T. Z., Meenen, B., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwackeleneh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, I., Pfankuch, C., Pioppert, F., Poindexter, A., Popovic, D., Primus, E., Pul, L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Saverly, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, I., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., Zvon

Niederhausem, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.
Unpublished
Rat Genome Sequencing Consortium.
2 (bases 1 to 183078)
Direct Submission
Submitted (03-OCT-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 183078)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 15, 2002 this sequence version replaced gi:23477788.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with 'N's to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KCMW
Center clone name: CH230-375A21
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 168021 bases at least Q40
Consensus quality: 170281 bases at least Q30
Consensus quality: 171978 bases at least Q20
Estimated insert size: 170592; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 180012: contig of 180012 bp in length
* 180013 180112: gap of unknown length
* 180113 181340: contig of 1228 bp in length
* 181341 181440: gap of unknown length
* 181441 183078: contig of 1638 bp in length.
Location/Qualifiers
1. 183078
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-375A21"
1. 1242
/note="wgs end_extension
clone_end:T7"
5362..6656
/note="wgs end_extension
clone_end:T7"
6707..7817

FEATURES
source
misc_feature
misc_feature
misc_feature

/note="wgs end_extension
clone_end:T7"
complement(8960..9793)
/note="clone_boundary
clone_end:T7"
site:
end sequence:RXAW1111J
26621..173267
/note="clone_boundary
clone_end:Sp5
site:
end sequence:BZ263204"
173661..175381
/note="wgs end_extension
clone_end:Sp6"
178477..180012
/note="wgs end_extension
clone_end:Sp6"

Query Match 30.4%; Score 692.504; DB 1; Length 183078;
Best Local Similarity 66.2%; Pred. No. 0.15;
Matches 1371; Conservative 0; Mismatches 585; Indels 115; Gaps 27;

Qy 270 TTCCCTCTCTTCCAAACACITCTATTCTTCTGATTTCTATCTTGGCTCATTTTAACTC 329
Db |||||
Qy 42317 TTCAATGAATCTAGAAATCTTTAAAGTTCTTTTATTCTGTCTTGATTCATTTTCAATC 42376
Db |||||
Qy 330 AGTAGTCAGTTGTTTGGTTTCCATAAGTTTGTAAAGTTTCTGTGTTTCTGTGTTGTTG 389
Db |||||
Qy 42377 AATAGAGAGTTGTTTACTTTTATGAGTTTGAAGATTCTTAATGTTCTGTCATTTG 42433
Db |||||
Qy 390 TTGTTATCTAGATTTAAGCTGTGCTGCTGATGACATAGACATAGATTAATTTCAATGTC 449
Db |||||
Qy 42434 TGAATATCCAGCTTTAATCCATGATGATTAACAGGATGACAGGAGTTTCTATTTTC 42492
Db |||||
Qy 450 TTTTATCTGTCGAGACTTGTCTTTGTTTGAATATGATCTTCAATTTTGAGAGTTTCA 509
Db |||||
Qy 42493 TTATATCTGTTAAACATGCTTCTCTGAGTATGCTCAATTTTAGAGGAGTTCTTA 42552
Db |||||
Qy 510 GGGTGTCTGAC - AAGAAGGTACAGTCTTT - GTGTTTGGTGAATAGTCTGTAATATC 565
Db |||||
Qy 42553 TGAGGCATACAGAAAGAGTTTATCTTTTGTGTTGAATGAATGTTCTGTAATATC 42612
Db |||||
Qy 566 TC-TAGTCCACTTGGTTTATGACATCAGTTAGTCTCCAGCATTTCTCTGTTGGTTTTT 624
Db |||||
Qy 42613 TCTTGGGTCCATTTGGTTTATAACATCTGTTAGTTCCATATTTCTCTGTTAGATTTA 42672
Db |||||
Qy 625 GTTGAGATGACCTAACTGTTGGAGAGAGTGGGTATTGAAGTAGCCCACTATCTGTGT 684
Db |||||
Qy 42673 TCTG-TATGACCTGCTTATTGATGAGAGTTGGGTAAATGAAGTCTCCCATTTATCTGT 42731
Db |||||
Qy 685 -GAGTCAATATGATTTTAGCTGTAGCTGTGCTTTTATGAACTTGGGTGGGTGACATTG 743
Db |||||
Qy 42732 GGGGGTCAATATGATTTAAGCTATAGTAGTATTTCTATTACAAATTTGGGTGCTCTT 42791
Db |||||
Qy 744 TGTGTTGTCATAGACATTAAGAAATTCGAATGCTCTTGGTGA - TTTTCTCTTTGATGC 802
Db |||||
Qy 42792 TGCTTGGCCGAAATATTAAGAAATTTGAATTCCTTCTTGGTTGAGTTTCTTGGATGA 42851
Db |||||
Qy 803 CTATGATGATTTCTTCCCAATCTCATCTGCTGTAGTTTGGGTAAAGTCTATTAGTC-AG 861
Db |||||
Qy 42852 ATATGGGGTGTCTTCCCTATCTCTTTTGATAATTTTGGTTGTAGAGATTTTCACTAG 42911
Db |||||
Qy 862 ATATTAATGACTGTATCGCTTGTCTTGGGCAATTTGCTTAGAATATCTTTTCCA 921
Db |||||
Qy 42912 ATATTAATGCTATCATAGCTTACTTCTTAGTCCATTTGCTTGGAAATATCTTTCCA 42971
Db |||||
Qy 922 TCCTTTTACTTCTAAGGTGATGCTATCCATGGTAGTTGTCTTTTGGATGACAGTGA 981
Db |||||
Qy 42972 ACCCTTTAC -CTTGGGGTAATGTCTGTCTTGATGTTTCAGATTGTTGCTCTTGTACAGTA 43030
Db |||||
Qy 982 GGATGGATCTGTTTTCATATCCATTCCTTACCAGATCTTTTCTTAGAGAAATTAAG 1041
Db |||||

Taylor, T., Thomas, N., Thomas, S., Tincey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, R., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Woodson, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, X., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

On Nov 19, 2002 this sequence version replaced gi:23809119. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GVMU

Center clone name: CH230-447P6

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 180908 bases at least Q40

Consensus quality: 182758 bases at least Q30

Consensus quality: 183743 bases at least Q20

Estimated insert size: 184876; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html). NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 198343: contig of 198343 bp in length

* 198344 198443: gap of unknown length

* 198444 199545: contig of 1102 bp in length.

Location/Qualifiers

1. .199545

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-447P6"

1. .1634

/note="wgs end_extension

clone_end:T7"

misc_feature

misc_feature

14358. .15643

/note="wgs end_extension

clone_end:T7"

misc_feature

17817. .18653

/note="clone_boundary

clone_end:T7

site:

end_sequence:BZ250252"

misc_feature

complement(197253. .197758)

/note="clone_boundary

clone_end:Sp6

site:

end_sequence:BZ250253"

Query Match

Best Local Similarity 30.4%; Score 692.504; DB 1; Length 199545;

Matches 1371; Conservative 0; Mismatches 585; Indels 115; Gaps 27;

Qy 270 TTCCCTCTCTCTCCAAACACTTCTATTCTTCTGATTTCTATCTTGGCTCATTTTAACTC 329

Db 102775 TTCAATGAATCTAGAAAATCTTTAAAGTCTTTTATCTCTGTGATTCATTTTCATTC 102716

Qy 330 AGTAGTCACTGTTTGGTTTCCATAAGTTGTAAAGTTTCTCTGTTCTGTTGTTGTTG 389

Db 102715 AATAGAGAGTGTGTTTACTTTTTTAAGATTTTGAAGATTTCTAAAGTTTCTGTCATTC 102659

Qy 390 TTGTATCTAGATTTAAGCTGTGCTGTCAGATAGGACATAGAGTATTATTTCAATTGTC 449

Db 102658 -TGAATCCAGCTTTAATCCATGATGGTATAACAGGATCGAGGAAGTATTCTTCTATTC 102600

Qy 450 TTTTATCTGCGAGACTTGCTTTGTTTGAATATGATATCAATTTTGGAGATTTTCA 509

Db 102599 TTATATCTGTTAAACACTTGCTTTCTGCTGAGTATGCTCAATTTTAGAGGAAGCTTA 102540

Qy 510 GGGTGTCTGAC - AAGAGGTACAGTCTTT - GTGTTTTGTGTAAGTAGCTGTAAATATC 565

Db 102539 TGAGGCATACAGAGAGAGTTTATCTTTTGTGTTGAATGAATGTTCTGTAAATATC 102480

Qy 566 TC-TAGGTCCACTTGTTTATGACATCAGTTAGCTCCAGATTTCTCTGTTCTGTTTCTT 624

Db 102479 TCTTGGTCCATTTGGTTTATAACATCTGTAGTCTTCCATATTCTCTGTTAGATTTTA 102420

Qy 625 GTTGAGATGACCTAACTGTTGGAGAGAAATGGGGTATTGAAGTAGCCCATCTATCTGTGT 684

Db 102419 TCTG-TATGACCTGCCCTATTGATGAGATTTGGTATGAACTCCCATATCATTTGCT 102361

Qy 685 -GAGGTCAATATGATGATTTTATGCTGTAGTCTGCTTTGTTTATGAACTTGGGTGACATTC 743

Db 102360 GGGGGTCAATATGATTTAAAGCTATAGTAGTATTCTATTACAAATTTGGGTGCTCTTG 102301

Qy 744 TGTTTGTGTGATAGACATTAAGAAATTCATGTCCTCTCTGGTGA-TTTTCTCTTTTCATGC 802

Db 102300 TGCTTCGGCCGAAAAATTAAGAAATTCCTTCTTGGTTGAGTTTTCTCTTGGATGA 102241

Qy 803 CTATGTAGTATTCTTCCCAATCTCATCTGCTTATGTTTGGGTTTAAAGTCTATTAGTCT-AG 861

Db 102240 ATATGGGGTCTTCTCCCTATCTCTTTTGAATAATTTTGGTTTGTAGACTATTTCCTAG 102181

Qy 862 ATATTAAATGACTGTATCGGCTTCTTCTTAGGGCCATTTGCTTAGAATATCTTTTCCA 921

Db 102180 ATATTAAATGTCTACATCAGCTTACTTCTTAGTCCATTTGCTTGAATATCTTTTCCA 102121

Qy 922 TCCTTTTACTCTAAGGTGATGCTATCCATGTTAGGTGTGCTTTTGTGATTCAGCAGTA 981

Db 102120 ACCCTTTAC-CTTGGGGTAAATGCTGCTCCCTGATGTTCAAGATTTGCTCTTGTACAGTA 102062

Qy 982 GGATGGATCTGTTTTCATATCCATTTCTGTTACCCAGTATCTTTTCTAGAGAAATTAAG 1041

Db 102061 GGAGGGTAGATGCTGT-TTTTCAATTTCTGTAGACCATGCTCTTTTATTGGGGAATTGAG 102003

Qy 1042 ATCATTGAGTCAATGATGTTGAGAAATATCAATGACAGCTGTTTGTGGGATCTTGTATC 1101

Db 102002 ACC-----ATATTGAGAGATATCGATGACACTGATTGTTGATTCTTGTATT 101955

FEATURES

source

QY 508 TAGGTGCTGACAGAAGGTACAGTCTTTGTTTGGTGAATAGTCTGTAAATATCTC 567
DB 59447 TCAGGTACTTAGAAGAAGGAATATCTTTTATGTTTGCTAAATGTTCTGTAGATAT -GG 59505
QY 568 TAGGTCCACTTGGTTTATGACATCAGTAGCTCAGCAATTTCTGTTTCTGTTTGGTT 627
DB 59506 TAGGTCCACTTGGTTTATGACATCAGTAGCTCAGCAATTTCTGTTTCTGTTTGGTT 59565
QY 628 GAGATGACTAACTGTTGGAGAGAAATGGGGTATTAAGATAGCCCACTATCTGTGTGAG 687
DB 59566 G-GATGACCTGTCCATTTGGTGAGAGTGGGGTGTGAACCTTTCCACATGCTAATGTGTAG 59624
QY 688 GTCAATATGATGATTTTACGTAGCTGTCTGTTTGTATTAAGAACTTTGGTGACATGTGTT 747
DB 59625 TTCAAAGTGAGATTTAAAGATCTTTTCTTTTGTGAGTGTGAGTGGCTTTGCAT 59684
QY 748 TGGTCCATAGACATTAAGAAATGCAATGCTCTCTGGTGGAT -TTTCTTTTGAAGCCCTAT 806
DB 59685 GTGGGCAATATCATTCAGAAATGAGATGCCATCTGTGGATGTTTCTTTGATGAATAT 59744
QY 807 GTAGTATTTCCCAATCTCATCTGCTTGTAGTTTGGGTTTAAGCTATATAG -TCAGATAT 865
DB 59745 AAGATATCTTCCCTATCTCTTTTGATTAATTTGGTTGAGATTTCTATTTTAAATAT 59804
QY 866 TAAATGACTGTATCGGCTTCTCTTAGGCCATTTAGCCATCTCTTAGATAT -CTTTTCCATCC 924
DB 59805 TAGAATGGCTGCTCCATCTGCTTCTGGTGTAGTGGCTTGGAAATGTTTTCATAT 59864
QY 925 TTTTACTCTAAGGTGATGCTATC ---CATGTAGTGTGCTTTTGGATGCGACGATA 981
DB 59865 CCTTACTCTGAGTAATGCTATCTTTTCTGAAGCCCTGTTCTCTGTATGACGAGAA 59924
QY 982 GATGGAATCTTTTTCATATCCATCTGTATCCAGTATCTTTTCTAGAGAAATTAAG 1041
DB 59925 TGATAGATCTTCTTTTGCATCCATCTGTTAGCCCTGTCTTTTATATGCAATATGAG 59984
QY 1042 ATCAATGAGTCAATGATGTGAGAAATTAATGACAGAGTGTCTGCAATCTCTGTTATC 1101
DB 59985 -----TCCATGATATTTGGAAATTAATGACAGAGTGTCTGCAATCTCTGTTAT 60036
QY 1102 TTGCACTTGTGAAGT 1161
DB 60037 TTG-----ATGT 60084
QY 1162 TGCTGT 1221
DB 60085 TGCTGT 60144
QY 1222 ATTTTGGCTGGAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1281
DB 60145 TTTTCTGGTGTGAATTTATTTATTTCTGGGTATTTATTTATTTATTTATTTATTTATTT 60204
QY 1282 TTGAAGTTTCTCCT -AGCCTTTCTTAGGTCTGCATTTGAAGATAGATATTTCTTACA 1339
DB 60205 TTGGAATTTCTCCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 60263
QY 1340 TCTGATTTTATCTTGAAGTCTTCTTCTTCTCCTCACTATTGTGACAGAAAGTTTCTTAA 1399
DB 60264 TTGATTTTGGATGAATTTCTGTTT-----GTATGGGATGATGTTTGTCTGT 60317
QY 1400 GTGCACTAGTCTGGCTGACATCTGTAGTCTCTGTGAGTCTGTAGACATCTGTGACAGG 1459
DB 60318 TTATATTAGTCTGGGGGGTTTCCATGCTCTCTGTAGTCAGAAAGACATCTGCCCAGGC 60377
QY 1460 CCTTCTTACATTTGAGTTTCTATTGGAAGTCAGGTGTAATTTCTATACATCTGCTT 1519
DB 60378 CATGCTGGTTTATGAGTCT -GATGAGAACTTGAATGTAATTTCTGTAGTCTGCTT 60434
QY 1520 TATATGTAATTTGGTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1579
DB 60435 TATATGTAATTTGGTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 60493
QY 1580 TTTAGTATTTGATTTATATGCACTGTGGGAGTCTTTTCCGGTCCATCTATTTGTT 1639

DB 60494 TCTAGTG-TTTGATTTATATGTGACAGGAGGATTTTCTTTT-TGGTCCAACTATTTGTT 60551
QY 1640 GTTTTGTATGCTTCTTGTATACCTTGAAGGATCTCTTTCTCAAGGTTAGGAAATTTTCT 1699
DB 60552 GTTCTATAGCTTCTTTATGTTTATAGGTATGTGCTCCCTTTTAGTTTAGGAAATTTTCT 60611
QY 1700 TTTTGGTCTTGTGAAATATTTTCCCTGCTTTTGAACCTGCTTCTTCCCTTCC--- 1755
DB 60612 TGTATGATTTTGTGAAATATTTTCTGGACCTTTTGAAGTGGGATCCCCCCCCCTTC 60671
QY 1756 -----TCTATTCTCTTGGTTTTCATAGTGTCTCTGGCTTCTTGGATGTTTAA 1804
DB 60672 TCTTCTATTTATTTAGGCTTCGTAATTTTCAAGTGTCTCCAAATTTCCAGATGTTTA 60731
QY 1805 TGCTCGATTTATTTAGACTTAACATTTTCTTTGACAAAGGATACCATTTCTTCTATCTT 1864
DB 60732 TGTAAAAAAGCTTCTATATTTGACATTTTATTTGACCAA-----TATATTGTCTATCCT 60785
QY 1865 GTCTTCACTGCTGAGATCTCTCTTCTATCTCTGTAATTTCTGCTAGTGAAGGCTTGTCTC 1924
DB 60786 CTCTTCTACGCCCTGAAATTTCTCTCATCTAGTCTGTAATTTCTGTAATGATCATGCTTGAGTC 60845
QY 1925 TGAGGTTCTCTGTTGGGTT-CTTAATTTTTCATTTCCAGATTTCTTCTAGTTTGGGTTT 1983
DB 60846 TGTAGTTCTCTGCTCTCTTACTAGGTTTTCCTCCTCAGGATTTCTCTAGTTTGTCTTT 60905
QY 1984 GTTTATTAATCTATTTTCCACTTTCAGTCTCAGGCTCGAAATGTTTACTCATTTCTCCTCC 2040
DB 60906 CTTTATTTGGTGTATTTGTTATTTTTCAGATCATGAACAGTTTATTCATCTCCTTACCTG 60965
QY 2041 AGTATTTTACATTTTCTAGGTTTCTTTAATGGAATTTATTCATTTCTCTTCAAGGACCTT 2100
DB 60966 TTTGTTGTAATTTCTGTAATTTCTTTTATGGAATTTATTAATTTCTTAAAGGACTC 61025
QY 2101 TTATGAATTCATAAAATGTAATTAAGTCTTGGCTTGTCTGCTTGTCTAGCTATGTTGATTC 2160
DB 61026 TATTGCTATATAAGATTTGATTTAAGGGCATCATCTAGGCTTCACTTACCTTAGGATG 61085
QY 2161 TCAGGGCC-----TATTGTAATAGGTTTATAGCAGGACATATTTGCTTCTGTTGTT 2211
DB 61086 TCTAGCACTTGTGTAGTAGATAGTGGGCTCTGGAGATGCCAATTTGCCCTGATTTGTT 61145
QY 2212 ATTCTGTGTGTTTGTCTTTGGCATATAGCGGCTGAG-TTTGGGATGATTTGAATTTCTA 2270
DB 61146 GTTG-ATTTTCTTACACTGGCTTTAGGCATCTGGGCTTTGTGTTATTAATCAGTTTA 61203
QY 2271 GGTCCTGAT 2279
DB 61204 GATGCTGAT 61212

RESULT 42

AC098032
LOCUS AC098032 225977 bp DNA linear HTG 10-MAY-2003
DEFINITION Rattus norvegicus clone CH230-185C21, *** SEQUENCING IN PROGRESS

ACCESSION AC098032
VERSION AC098032.6 GI:30521836
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 225877)
Murny,D,Marle, Metzker,M, Lee, Abramzon,S, Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Cesar,H., Center,A.,

Qy	1065	AATTATCAATGACGAGTGTGTGGAGATCTTGTTATCTTGGCACTCTGTGGAAGTGCTGTGT	1124
Db	74508	AGTTATCAATGACGAGTGTGTGGAGATCTTGTTATCTTGGCACTCTGTGGAAGTGCTGTGT	74454
Qy	1125	GT	1184
Db	74453	TGTTGGTAGT	74394
Qy	1185	CTGTGTCTTGT	1243
Db	74393	TGT	74334
Qy	1244	TATTCATATATTTCTTGAATGCGGTAACATCTTTAGATTGAAGTTTTCCTCTAG--CC	1301
Db	74333	CTCT--GTACTTCTTGGATATAAATTGACCTCTTGAGTTTGAAGATTGTCTTCTAGTCC	74276
Qy	1302	TCTTTAGGTCTGCATTTGAAGATAGATATCTTTACATCTCAATTTTATCTTGAAGTCT	1361
Db	74275	TATTTAGGACTGGATTTGTAGATAGATAAAACCTAAATTTGATTTTATGTGAAA----	74220
Qy	1362	TTTCTTCTCCAACTATGTGACAGAAAGTTTTCCTAAGTGCAGTAGTCTGCCTGACAT	1421
Db	74219	TTTTTTCTCCATTTATGTGGCTGAAAGTCT-----GAGTCTAGTGGGAT	74172
Qy	1422	CTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGCAGGGCTCTTCTTACATTTTGAATTTCT	1481
Db	74171	CTGCAGACTCTTAGAATTCATAGAACATCTGTCCAGGTCCTCTGCTTTTACAGTCTCC	74112
Qy	1482	ATTGGAAAAGTCAGGTGTATTTCTAATACATCTGCCCTTTATATGTTAAATGGTCTTTTT	1541
Db	74111	ACTGTAAAA--CAGATATTAATCTATTAGCGGTCTCTATATGCTACTTGTGCTTTTCC	74054
Qy	1542	CCCTTGACATCTTTAATATCTTTCTTGTCTATACTTTTAGTGATTTGATTTATATGC	1601
Db	74053	CCCTTATATCTTACAAATATCTTCTTTGTT--TGATGTTTATTTGTTTGAATTTATATGT	73995
Qy	1602	ACTGTGGGAGTTCTTTCCGGTCCAACTATTTGGTGTTTT--GTATGCTTCTTGTA	1658
Db	73994	ACTGAGGCGAGTTCTTTCTGTCTCTGTTTATTTTGTATTTCTGAAGTAAATCTCTGTA	73935
Qy	1659	CTTTGATAGGCATCTCTTCTCAAGGTTAGGAAATTTTCTTTTGGTTTTCTTGAAAA	1718
Db	73934	CTTTGATAGACATCTCTCTTTAGATTAGAAATTTCTCTCTGTGATTTCTATTGAAAA	73875
Qy	1719	TATTTTCCCTGTTTTGACCT--GCCTTCTTCCCTCTCTATCTCT-----T	1765
Db	73874	TGTTATCTGTGCCCTTGATCTGGGTTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	73815
Qy	1766	TGGTTTTTGATAGTGTCTCTGGCTTCTCGATGTTTTTATGCTCTGATTTATTTAGACTT	1825
Db	73814	TGGTCTTTTCACTCTGTTTCTGATTTCTGAACTCACTTTGTAGCTGAAGCTTTTGTAGTTT	73755
Qy	1826	AACATTTTCTTTGACCAGGATTCATTTCTTCTATCTTGTCTTCTACTGCTGAGATCT	1885
Db	73754	AACATCTATTGATCTGATAAGTCCATTTCTTCTAGCATGTCTTCAATGCTTGTGATCT	73695
Qy	1886	CTCTCTATCCTGTGATTTCTGTGACGAGGCTGTCTGTGAGGTTCCGTGGGTTCTT	1945
Db	73694	TTCCTGTGCGCTGTATCTTTTGGTGAATCTTGCCCTATCAGGCTCCCTTTCAAGTTCT	73635
Qy	1946	AAATTTTTCATTTCCAGATTTCTCTCAGTTTGGGTTTTGTTTTTAATTTCTATTTCCACT	2005
Db	73634	AAATTTTTCATTTCCAGATTTACCCTCACTTTGGATTTT--CCTATTGACTTATTTCCATT	73576
Qy	2006	TTGAGGTCCTGAAATGTTTTTACTATTTTCTCTCCAGATTTTACATTTTTCATAGGTTCT	2065
Db	73575	TTGAGGTCCTGAACTGTGTTTTATCCA--TTCCTCTCAAAGGCTATGTTTTCAITGATTTCT	73517
Qy	2066	TTAATGGATTTTATTCATT	2083
Db	73516	TTAAGAGACTTTTGTCIT	73499

RESULT 44
AC122982
LOCUS
DEFINITION

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

TITLE
JOURNAL
REFERENCES
AUTHORS
TITLE
JOURNAL

REFERENCE

AC122982 261378 bp DNA linear HTG 06-NOV-2002
Rattus norvegicus clone CH230-341M21, *** SEQUENCING IN PROGRESS
***.

AC122982
AC122982.4 GI:24635417
HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED.
Rattus norvegicus (Norway rat)

Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

1 (bases 1 to 261378)

Muzny, D. Marie, Metzker, M. Lee, Abramzon, S., Adams, C., Alder, J. J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, K., Barnstead, M., Benahmed, F., Biswal, D., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, J., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Gear, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. I., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idiebor, D., Johnson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshuwa, L., Lohsedeg, H., Lozada, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Manwhiney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokemele, C., Okunou, G., Olareunsgaon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shwartsbeyn, A., Sisson, I., Sitter, C. D., Smajda, D., Sneed, A., Sodergren, E., Song, X.-Z., Sotelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tinney, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlarczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinsteins, G. and Gibbs, R. A.

WELLS LOCK, G.: and
Direct Submission

Unpublished

2 (bases 1 to 261378)

Worley, K. C.

Direct Submission

Submitted (26-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 261378)

AUTHORS Rat Genome Sequencing Consortium.
 TITLE Direct Submission
 JOURNAL Submitted (06-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 COMMENT On Nov 6, 2002 this sequence version replaced gi:23907833. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu

----- Project Information
 Center project name: GYOK
 Center clone name: CH230-341M21
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 195499 bases at least Q40
 Consensus quality: 197131 bases at least Q30
 Consensus quality: 198079 bases at least Q20
 Estimated insert size: 203971; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 1 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 261378: contig of 261378 bp in length.
 Location/Qualifiers
 1..261378
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-341M21"
 1..1697
 /note="wgs end_extension
 clone_end:17"
 53300..55403
 /note="wgs end_extension
 clone_end:17"
 59206
 /note="clone_boundary
 clone_end:17
 site:
 end_sequence:B2188858"
 complement(258171..258994)
 /note="clone_boundary
 clone_end:Sp5
 site:
 end_sequence:B2188859"

misc_feature
 misc_feature
 misc_feature
 misc_feature
 Query Match 29.9%; Score 680.301; DB 1; Length 261378;
 Best Local Similarity 69.2%; Pred No. 0.23;
 Matches 1258; Conservative 0; Mismatches 477; Indels 83; Gaps 25;

Qy 306 TCTATCTTGGCTCATTTTAACTCAGTAGTAGAGTTGTTGGTTTCCATAAGTTTGAAGT 365
 Db |||||
 Db 95927 TTTGTCCTGAGCCATTTTTCATTCACCTAGTAGTCTCTGCTCTCCATGCATTCACAAC 95986
 Qy 366 TTTCTGTGTTTCTGTGTTGTTGTTTATCTATCTAGATTTAAAGCTGTGGTGGTCAGATAGG 425
 Db |||||
 Db 95987 TTTCTGTGTTTCTGTGATA---TTGTTGGTATCCAGCTTTAATCTGTGGTGGTCAGATAGG 96043
 Qy 426 ACATAGAGTATTTTCAATTTGTTTATCTCTGTGCAGACTTCTTTGTTTGAATATG 485
 Db |||||
 Db 96044 ACATAAGTATTTTTCGGTCTCTTCATCTGTTGAGACTTTCTTTGTGTCAGATATG 96103
 Qy 486 TATTCAATTTTGG--AGAGTTTTCATAGGTGCTGCACAAGAAGTACAGCTTTTGTGTTT 543
 Db |||||
 Db 96104 TGGTAACTCTGGAAAAGTTCATCGGTTCTTGAGAAAAGGTATATTTCTTTGTGTTG 96163
 Qy 544 GTGAAATAGTCTGTAAATATCTCT--AGGTCCACTTGGTTTATGACATCAGTTAGTCCA 602
 Db |||||
 Db 96164 ATTGAAATGTTTGTAAATACATGTTAGGTCTGTGTGGTTTATGACATTAGATATCTCCA 96223
 Qy 603 GCATTTCTCTGTTTCTGTTTCTGTTTTCAGATGACTTAACCTGTTGGAGAGAAATGGGTATG 662
 Db |||||
 Db 96224 GCATTTCTCTGTTAG--TTTTTGTGTAGGTGACCTGCCCTATTGGTGAGAGAGGTATTA 96282
 Qy 663 AAGTAGCCACTATCTGTGTGTGA--GGTCAATGATGATTTTAGCTGTAGCTGTGCTTGT 721
 Db |||||
 Db 96283 ATGCTCCCACTATCAGTGTACGAGGTCAATTATATGAGGTGGCTTTAGTATGTTCT 96342
 Qy 722 TTTATGAATCTGGTGCACATTTGTTTGGTGCATAGACATTAAGATTCGAATGCTCTCT 781
 Db |||||
 Db 96343 TTTATGAATCTGGTG--CCCTGTGTTTGGTGATAAGTACTAAAATTTCAATATCCCTC 96401
 Qy 782 TGTGTG-----ATTTTCTCTTGATGCCCTATGTAGTATTCTTCCCAATCTCATCTGCT 834
 Db |||||
 Db 96402 TCTTAGTGTGTTTTTCTCTCTTTTAAATATGATGAGTCTCTCCCTAGCTCTCTGATT 96461
 Qy 835 AGTTTGGGTTTAAAGCTA--TTAGTCAGATATTAATGACATGATCGCTGCTCTCTTA 893
 Db |||||
 Db 96462 AGTTTGTATGATGAAGCTATTTTGTAGATATTAATAAGCTATTCTGGCTTGTCTTGT 96521
 Qy 894 GGCCCAATTTGCTTAGAATATCTT-----TTCCATCTCTTTACTCTAAGGTGATGCTAT 947
 Db |||||
 Db 96522 GTGCCATTCATTAATAATATCATTTGCCATCTTCCCTGCCCCCAAGATGATGCTAT 96581
 Qy 948 CC---ATGCTAGTCTCTTTTGTGATGAGCAGTAGAGTGAATCTTGTTCATATCC 1004
 Db |||||
 Db 96582 CCTTAATGTATAGATGTGTTCTTAGATGACAGAAAGATGAA--TAGTTTTTCATATTC 96639
 Qy 1005 ATTCTGTATCCCAAGTATCTTTTCTAGAGAAATTAAGATCATTTAGTCAATGATGTTGAG 1064
 Db |||||
 Db 96640 ATTCGGTTTAATAGTCTGTTGTAGGGCAGTTGA-----AACCATTCACATGAG 96691
 Qy 1065 AATTATCAATGACAGCTGTTTGTGGATCTTGTATCTTTGCACTGTGTGAAGTGTGTGT 1124
 Db |||||
 Db 96692 AGTTATCAATGACCACTATTGTTGATTCAGTATTTTGT-----TTGTGTGGTGAATGA 96746
 Qy 1125 GT 1184
 Db |||||
 Db 96747 TGT 96806
 Qy 1185 CTGTGTCTTGT 1243
 Db |||||
 Db 96807 TGT 96866
 Qy 1244 TTTATCATATTTTCTTGAATGTGGGTAAACATCTTTAGATTTGAAGTTTCTCTCTAG--CC 1301
 Db |||||
 Db 96867 CTCCT--GTACTTTCTTGGGATATAATTGACCTCTTGAGTTTGAAGATTTGTTTCTAGTGCC 96924
 Qy 1302 TCTTTAGTCTGTGCATTTGAAGATAGATATTTCTTACATCTCTGATTTTATCTTAGAATGTC 1361
 Db |||||
 Db 96925 TATTATAGACTGGATTTCTAGATAGATAAAACCTAAATTTGATTTTATTTGTGAAA--- 96980
 Qy 1362 TTTCTTTTCCAACTATTGTGTGACAAAAGTTTTTCTTAAGTCGATGCTGCGCTGACAT 1421

Db 96981 TTTTTCCTCCATTTATTGCTGGCTAAAGTCT-----GAGTCTAGGTGGGAT 97028
 QY 1422 CTGTAGTCTCTGAGTCTGTAGCACATCTGTGAGGCGCTCTTACATTTGAGTTCT 1481
 Db 97029 CTGAGAGCTCTTAGAATTCATAGAACATCTGTCCAGGTTCTCTGGCTTTTACAGTCTCC 97088
 QY 1482 ATTGAAAGTCAGGTGAATCTTAATACATCTGCCCTTATATGTTAATGTCGCTTTT 1541
 Db 97089 ACTGTAAAA--CAGATATATTCTATTAGGCGTCTCTATATGCTACTTGGCTTTTCC 97146
 QY 1542 CCCGTGATCTTTAAATCTTTCTTTGTTCTATACCTTTTGTAGTGAATTTATATGTC 1601
 Db 97147 CCCGTATATCTACATATCTTTCTTTGTTT-TGTATGTTTATTGTTGATTAATGT 97205
 QY 1602 ACTGTGGGAGTCTTTTTCGGTCCAAATCATTTGTTGGTGT-----GTATGCTCTTTGTA 1658
 Db 97206 ACTGAGGCGAGTCTTTTCTGTTCTGTTCTTTATTGATGTTCTGAAGTAAATCTTTGTA 97265
 QY 1659 CTTGTAGAGGATCTCTTCTCAAGTTAGAAATTTTCTTTTGTGTTTCTTGAAAA 1718
 Db 97266 CTTGTATAGACATCTCTCTTTCTTAGATTAGAAATTTCTCTGTTCTATTTGAAA 97325
 QY 1719 TATTTTCCCTGCTTTTGACCT-GCCTTCTTCCCTCTCTTATCTCT-----T 1765
 Db 97326 TGTATCTGTGCTTTGATCTGGGTTCTTCTCTCTCTCTATCTCTACTAGTCTTAGAT 97385
 QY 1766 TGGTTTTGCATAGTCTGTGCTCTGCTCTGATGTTTATGCTCGGATTTATTAGACTT 1825
 Db 97386 TTGTCCTTCCACCTGTTTCAGATTCTGTAACACTTTGTAGCTGAAGCTTTTAGATTT 97445
 QY 1826 AACATTTTCTTTGACCAAGTATCATCTTCTTCTTCTGCTTCTTCTCACTGCTGAGATCT 1885
 Db 97446 AACATCTCATTTGACTGATAGTCACTTTCTTCTACCACTGCTTCAATGCTTGATCT 97505
 QY 1886 CTCCTCTATCTCTGATCTGTCTAGTCAAGCTTGTCTCTGAGTCTCTGTTGGTCTTT 1945
 Db 97506 TTTCTCTGTGCTTGTATTCTTTGTTGTAACCTTGCTATCAGGCTCCCTTTCAAGTCTCT 97565
 QY 1946 AATTTTTCATTTCCAGATTTCTTCAAGTTGTTGTTGTTTATTAATCTTATTTCCACT 2005
 Db 97566 AATTTTTCATTTCCAGATTTCCCTTCAAGTCTTGTGATTTT-CTTATGACTCTATTTCCAT 97624
 QY 2006 TTCAGTCTGAAATGTTTACTCATTTTCTTCCAGTATTTACATTTTCAAGTTTCT 2065
 Db 97625 TTCAGTCTGAACTGTTTATTTCA-TTCCCTTCAAGGCTATGTTTCTATTTCTTCT 97683
 QY 2066 TTAATGGATTTATTCATT 2083
 Db 97684 TTAAGAGACTTTTGCTTT 97701

 RESULT 45
 AC093959
 LOCUS Rattus norvegicus clone CH230-3B9, WORKING DRAFT SEQUENCE, 2
 DEFINITION
 AC093959
 ACCESSION
 VERSION AC093959.7 GI:30467773
 KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 Rattus.
 1 (bases 1 to 274459)
 REFERENCE
 AUTHORS Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alshrocks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,

Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escoto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gebregeorgis, E., Geet, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, J.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowis, C., Kraft, C.J., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorensuhen, L., Loulseghe, H., Lozano, R.J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangun, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mawhiney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K.,
 Nwankwelen, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plummer, F., Polindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
 Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 Riley, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Shen, H.,
 Sanders, W., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbey, A., Sisson, I., Sitter, C.D., Smajs, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstock, G. and Gibbs, R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 274459)
 Worley, K.C.
 Direct Submission
 Submitted (13-SEP-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 274459)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (09-MAY-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On May 9, 2003 this sequence version replaced gi:24819560.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>

Db 135067 TTCTATTATCTAGGCTTAATCTTTTATGGTTTCCCAAAATTCCTGAATGTTTGTGT 135126
QY 1808 CTGGATTATTTAGACTTAACATTTCTTTGACCAAGGTATCCATTCTTCTATCTTGTGTC 1867
Db 135127 CAGATCTTTTAAATTTACATTTTCTTAACTGATGATCAACTTCTTCTACTGAAT 135186
QY 1868 TTCACCTGCTGAGATCTCTCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTGA 1927
Db 135187 TTCTATGCTGAGATCT 135246
QY 1928 GGTCTCTGTTGGTCTCT-TAATTTTCTCAATTTCCAGATTTCCCTCAGTTTGGGTTTGT 1986
Db 135247 TGAATCTGTTGCTTAATCTAGATTTTCTAATCTCCAGGATTCCTCAGTTTGTGTTTCTT 135306
QY 1987 TATTAATCTATTTCCACTTTCCAGTCTCTGAAATGTTTACTCATTTTCTCCCAATTT 2046
Db 135307 TACTGCTTCTAATTTTCATTTTCAATTAACAGTTTATTTATTTTC--CTTCAACCAATTGAT 135364
QY 2047 TACATTTTCATAGTTTCTTTAATGGAATTTATCTATTTCTCTCTTCAAGGACCTTTTATGA 2106
Db 135365 GACATTCCTCTGTGATTTCTTTAAGAAATCCATTCATTTCTCTCTTTTAAAGAGGAT----- 135418
QY 2107 ATTCATAAAATGATGTTAAGTTCCTTGGCTTGTGCTTCCAGCTATGTTGCATTTCT----- 2161
Db 135419 -----GTGAGTCAATGCTCTTATGCTTCACTTGTGCTAAATATCTAGG 135462
QY 2162 ----CAGGCGCTATGCTAATAGGTTTATAGCGGACATATTTGCTCTGTTGTTTATGTC 2217
Db 135463 GTTCTTGGAGTACTGTAATCACTCTCTGTTGGTGCATACTGCTCTGTTGTTGTTGTTGAT 135522
QY 2218 TGTGTTTTT 2226
Db 135523 TGTGTTCTT 135531

Search completed: August 24, 2004, 16:54:55
Job time : 11290 secs

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2004, 06:30:30 ; Search time 2253 Seconds
(without alignments)
3.614 Million cell updates/sec

Title: US-10-664-775-4

Perfect score: 2279

Sequence: 1 gatcaactctctagtgaag.....ttgtaattcttaggtgctgat 2279

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 0.5

Searched: 39 seqs, 1786346 residues

Total number of hits satisfying chosen parameters: 78

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : rng4.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	591.499	26.0	19619	1 ABK47192	Mouse Contig 1A co
C 2	588.5	25.8	44576	1 AAL61522	Cosmid CVO14 conta
C 3	588.3	25.8	29604	1 AAX83005	Partial mouse WRN
C 4	585.1	25.7	173810	1 ABN85752	Mouse chromosome 1
C 5	578.299	25.4	173810	1 ABN85752	Mouse chromosome 1
C 6	578.2	25.4	8048	1 ABN85760	Toxicity-related g
C 7	578.2	25.4	8048	1 AD852444	Primary rat hepat
C 8	565.399	24.8	51259	1 AAX83007	Partial mouse WRN
C 9	555.699	24.4	47115	1 ADA02627	Mouse Flt3 carcino
C 10	555.699	24.4	47115	1 ADA02627	Mouse Flt3 gene.
C 11	525.898	23.1	96597	1 ADA02501	Mouse Bach2 gene.
C 12	525.898	23.1	96597	1 AD872239	Mouse Bach2 gene.
C 13	521.797	22.9	51259	1 AAX83007	Partial mouse WRN
C 14	486.098	21.3	9320	1 AAD51996	Mouse Cyp3A11 gene
C 15	486.098	21.3	12275	1 AAD51996	Mouse Cyp3A11 prom
C 16	475.598	20.9	3449	1 AAF92531	Rat T2804 nucleoti
C 17	453.098	19.9	96594	1 ADC85257	Mouse Ptpkr genom
C 18	453.098	19.9	96595	1 ADA02777	Mouse Ptpkr gene.
C 19	453.098	19.9	96595	1 AD872515	Mouse Ptpkr gene.
C 20	381.798	16.8	41637	1 AD802837	Mouse Map3k8 carci
C 21	381.798	16.8	41637	1 AD872575	Mouse Map3k8 gene.
C 22	381.798	16.8	41637	1 ADC85316	Human Mef2c coding
C 23	328.699	14.4	2920	1 AAX35043	Murine cDNA isolat
C 24	328.699	14.2	16442	1 AAX83006	Partial mouse WRN
C 25	302.299	13.3	41400	1 AAX77189	Mouse BAC279 clone
C 26	288.196	12.6	46765	1 AAX99306	DNA encoding Aldeh
C 27	280.299	12.3	1632	1 AD859208	Toxicity-related g
C 28	278.099	12.2	3903	1 AD859149	Toxicity-related g
C 29	277.897	12.2	70665	1 ABT10716	Human breast cance
C 30	269.097	11.8	56583	1 AAX35003	Human adenosine re
C 31	269.097	11.8	56583	1 AAF21125	Human low adenosin
C 32	269.097	11.8	56583	1 ABX96819	Human nucleic acid
C 33	268.097	11.8	50000	1 AAG64139	Nucleotide sequenc

C 34	267.297	11.7	136284	1 ABK83575	Human cDNA differe
C 35	256.797	11.3	32167	1 AAL05509	Human reproductive
C 36	251.899	11.1	10409	1 AAV42558	Mouse dectin-2 gen
C 37	249.899	11.0	2396	1 ABV76780	Mouse tyrosinase g
C 38	249.899	11.0	3649	1 ABV76778	Mouse tyrosinase g
C 39	249.899	11.0	3711	1 ABV76777	Mouse tyrosinase g
C 40	249.898	11.0	125910	1 AAC64370	Human KCNQ5 (KCN6q
C 41	238.198	10.5	6207	1 AAS63340	Chemically pretrea
C 42	149.2	6.5	70665	1 ABT10716	Human breast cance
C 43	141.4	6.2	125910	1 AAC64370	Human KCNQ5 (KCN6q
C 44	71.4	3.1	96594	1 ADC85257	Mouse Ptpkr genom
C 45	71.4	3.1	96595	1 ADA02777	Mouse Ptpkr carcin

ALIGNMENTS

RESULT 1

ABK47192/C
ID ABK47192 standard; DNA; 19619 BP.

AC ABK47192;

DT 05-JUN-2002 (first entry)

XX Mouse Contig 1A containing a fragment of the Ots1-B7 gene.

XX ds; nucleic acid library; immune response; asthma; DC-SIGN;

KW airway hyperresponsiveness; bronchoalveolar manifestation;

KW signature sequence; SS; chronic obstructive pulmonary disease; COPD;

KW allergic disease; rhinitis; atopic dermatitis; urticaria;

KW autoimmune disease; multiple sclerosis; inflammatory bowel disease;

KW allograft rejection; infectious disease; Ots1-B7.

XX Mus sp.

XX WO200214366-A2.

XX 21-FEB-2002.

PF 16-AUG-2001; 2001WO-NL000610.

PR 16-AUG-2000; 2000EP-00202867.

PA (UYUT-) RIJKSUNIV UTRECHT.

PI Groot PC, Van Bergenhenegouwen BJ, Van Oosterhout AJM;

DR WPI; 2002-241888/29.

XX P-PSDB; AAU88024.

XX Nucleic acid library comprising genes which are capable of initiation, progression and suppression of an immune response, especially an immune response observed with airway hyper-responsiveness of asthma.

PS Example 10; Fig 9; 120pp; English.

XX The invention relates to a nucleic acid library comprising genes or their fragments which are capable of modulating an immune response observed with airway hyperresponsiveness and/or bronchoalveolar manifestations of asthma. Also included are a method for modulating an immune response of an individual comprising modulating a gene comprising a nucleic acid at least functionally equivalent to a nucleic acid identifiable by a signature sequence (SS) given in the specification such as R1-SO-R1-A11, StOI-A10, SvO2-1-C11, StOI-A12, and R1-SO-R1-B7, a substance (for use as a medicament) capable of modulating a gene comprising a nucleic acid at least functionally equivalent to a nucleic acid identifiable by SS and the use of a proteinaceous substance derived from a nucleic acid at least functionally equivalent to a nucleic acid identifiable by SS for the production of an antagonist (for use as a medicament) against the substance. The antagonist and substance are useful for the treatment of an immune response observed with airway hyperresponsiveness and/or bronchoalveolar manifestations of asthma. The method is useful for


```
QY 1659 CCTGTAGGACATCTCTTCTCAAGGTAGGAAATTTTCTTTTGGTTTCTTTGAAAA 1718
Db 15128 TGCCTATGGGTATCTCTTTTGTAGGTAGGAGTTTCTCTATGATTTTGTGAAGA 15069
QY 1719 TATTTTCCCTGCTTTGACCTG-----CCTTCTCCCTTCTCTATTTCT 1764
Db 15068 CATTTACTGCTCTCTTTGAGCTGGAGTCTTCACTCTCTCTATACCTATTTACTCTTAGGT 15009
QY 1765 TTGGTTTTCATAGTCTCTGGCTCTCTGCTGATGTTTATGCTGGATTTATTTTAGACT 1824
Db 15008 TTGATCTTCTCATGTAGTCTGATTTCTGTAAGTTTGGACAGTAGCTTTTCCGCT 14949
QY 1825 TAAATTTTCTTTTGACCAAGGTATCCATTTCTTCTATCTTGTCTTCACTGCTGAGATTC 1884
Db 14948 TTACATTAATCTTTGACAGTTGAGTCAATGATTTCTATGGAATCTTCTGCTCTTGAGATTC 14889
QY 1885 TCTCTTCTATCTCTGATTTCTGTCAGTGAAGCTTGTCTCTGAGTTCC-TGTTGGGTTTC 1943
Db 14888 TCTCTTCCATCTCTTGTATCTGTTGGTGTATCTGATCTGATCTACGGCTCTTGTCTCTTCC 14829
QY 1944 TTAAATTTTTCATTTCCAGATTTCCCTTCAGTTTGGGTTTGTGTTTATTAATTTCTATTCCA 2003
Db 14828 TTGGTTTCTATATCCAGGTTGTTTCCATGTTCTTCTTGTATGCTTCTATTCCA 14769
QY 2004 CTTTCAGGTCCTGAAGTTTACTCATTTTCTCCAGTATTTACATTTTCTCATAGGTTT 2063
Db 14768 TTTTAAATCTTCAACTGTT-----TGATTGTGTTTCTCTGGAAT 14727
QY 2064 CTTTATGATTTTATTCATTTCTCTCTCAAGGACCTTTTATGAAATTCATAAATGTATGT 2123
Db 14726 CTTTCAGGATTTTCCGATTTCTCTCTGTAGCTTCTACTTGTCTCTAGGAGTCTCT 14667
QY 2124 TAAAGTCTTT 2133
Db 14666 TCACGTCCTT 14657
```

RESULT 3

AAx83005
ID AAX83005 standard; DNA; 29604 BP.

AC AAX83005;

DT 31-AUG-1999 (first entry)

DE Partial mouse WRN genomic sequence #1.

XX Mouse; WRN; Werner's syndrome; detection; diagnosis; autosomal;

XX recessive disorder; phenotype; ss.

XX Mus_musculus.

XX WO9724435-A1.

XX 10-JUL-1997.

XX 30-DEC-1996; 96WO-US020785.

XX 29-DEC-1995; 95US-0009409P.

XX 29-DEC-1995; 95US-00580539.

XX 30-JAN-1996; 96US-0010835P.

XX 30-JAN-1996; 96US-00594242.

XX 12-APR-1996; 96US-00632175.

XX (DARW-) DARWIN MOLECULAR CORP.

XX Oshima J, Fu Y, Yu C, Mulligan J, Schellenberg GD;

XX WPI; 1997-363671/33.

XX Isolated nucleic acid molecule encoding the WRN gene product - useful for

XX detection and treatment of Werner's syndrome, and related diseases.

PS Claim 1; Fig 7; 153pp; English.

XX This sequence represents a fragment of the genomic sequence containing
CC the coding region for the mouse WRN gene (AAX83004). The corresponding
CC human gene (AAX83001) encodes a protein related to Werner's syndrome. The
CC products can be used for the detection and treatment of Werner's syndrome
CC (WS), an autosomal recessive disorder with a complex phenotype, as well
CC as related diseases

XX Sequence 29604 BP; 7634 A; 5861 C; 5985 G; 10123 T; 0 U; 1 Other;

QY Query Match 25.8%; Score 588.3; DB 1; Length 29604;

XX Best Local Similarity 62.9%; Pred. No. 1.4;

XX Matches 1170; Conservative 0; Mismatches 567; Indels 123; Gaps 18;

QY 292 TCTATTTCTTGATTTCTATCTTGGCTCATTTTAACTCAGTAGTGAAGTTTGGTTTCC 351

Db 23722 TTCTTTTCTTATCTCTTCCCTTGACCAAGTATCATTTGAGAGTGTATTTCAGTTTCC 23781

QY 352 ATAAAGTTTGAAGTTTCTCTGTTTCTGTTGTTGTTGTTGTTTATCTAGATTTAAAGTGT 411

Db 23782 ACGTGAATGTTGGCTTTCCATTTATGTTGTTTATTAAGATC---AGCCTTAGGCCAT 23838

QY 412 GGTGCTCAGATAGACATAGATGATTTTCAATTTGCTTTTATCTCTCGAGACTTGTCT 471

Db 23839 GGTGCTCTGATAGATACATGGGACAAATTTCAATATTTTGTATCTATTTGAGCCCTGTTT 23898

QY 472 TGTTTTCATAGATGATTTTCAATTTTGGAGA--GTTTCATAGGGTGTCTGACAAAGGTATC 529

Db 23899 TGTGACCAATATATGTTGTCATTTTGGAGAAGTCCCGTGAGGTGTCTGAGAAGGTAT 23958

QY 530 AGTC-TTTGTGTTTGTGTAATAGTCTGTAATATCTCT-AGTTCACCTTGGTTTATGA 597

Db 23959 ATCCTTTTGTGTTAGGATAAAATGTTCTGTAGATATCTGTGAGGTCAATTTGTTTCATAA 24018

QY 588 CATCAGTTAGTCTCAGCAATTTCTGTTTGGTGTGTTTGTGATGATGACCTAACTGTTGA 647

Db 24019 CTTCGTGTTAGTTCACGTGTCCTCTGTTAG-TTCTGTTTCCAGCATCTGCTCTTGA 24077

QY 648 GAGAAATGGGTATTAAGTAGCCACATCTGTGTGTGAGGT-CAATATGTGATTTTACG 706

Db 24078 GAAAGTGTGTGTTGAAGTCTCCCACTATTATTGTTGTGAGGTGCAATGTATGTTGAGC 24137

QY 707 TGTAGCTGTCTCTTTTATGAATGGGTGACATTTGTTGTTGTCATAGACATTAAGA 766

Db 24138 TTTACTAAAGTGTCTCTAAATGAATGTGGCTGCCCTGTCATTTGTCGCTAGATATTCAGA 24197

QY 767 ATTGCAATGTCCTTGGTGGATTTT-CCTTTGATGCTTAATGATGATTTCTTCCCAATCT 825

Db 24198 ATTGAGTGTTCCTCTTGGAGGATTTTACCTTTGATGATGAAGTGTCCCTCTCTGTCT 24257

QY 826 CATCTGCTTAGTTTGGTTTAACTCTA-TTACTCAGATATTAAATGACTGATCTCGCT 884

Db 24258 TTTTGTATTAATTTGGGTTGGAAGTCAATTTTATCCGATTAATAATGGCTATCCAGCT 24317

QY 885 TGCTTCTTAGGGCAATTTGCTTAGATA-TCTTTTCCATCCTTTTACTCTTAAGGTGATGT 943

Db 24318 TGTTCCTTCAGTCCATTTGCTTGAAATTTGTTTCCAGCCTTTTACTCTGAGGTAGTGT 24377

QY 944 CTATCCAT---GGTAGGTGTCTTTTGGATGACAGAGTAGGATGAATCTGTTTCAT 1000

Db 24378 CTGTCTTTTCCCTGAGATGGGTTTCCCTGTAAGCAGCAGAAATGTTGGTCTCTTGTGT 24437

QY 1001 ATCCATTTCTGTACCCAGTATCTTTTCTAGAAATTAAGATCAATGAGTCAATGATGT 1060

Db 24438 AGCCAGTCTGTAGTCTATGTCCTTTTATTTGGGAATTTGAG-----TCCATGATAT 24489

QY 1061 TGAGAAATTAATGACAGTGTGTTGGATCTTGTATCTTGTGCACTTGTGAAGTGTGT 1120

Db 24490 TAAGAGATATTAGGAAAGTAATTTGTTGCTTCTCTTTTATTTTGTGTAGAGTTGGCA 24549

QY 1121 GT 1180

Db 24550 TTCGTTCTTGGGCTTCTCTCTTTGGTTG----- 24582
Qy 1181 GTGCTGTGTCTTGCTGTCTGTCTCTCCCTCTTTGATTTTGGCTGGAATTAT 1240
Db 24583 -----TTCAATGAT 24591
Qy 1241 TTATTAATCATATTTCTTGAATGGGTAACTTTTAGATTGAAGTTTTCCTCCTAGC 1300
Db 24592 TACTTTCTGGTGTCTTAGGGCGTGAATTCGGTCTTGTATTCCTTCTTCTGTTATT 24651
Qy 1301 CTTCTTT--AGGCTCTGCAATTTGAAGATAGATATTTTACATCTGATTTTATCTTAGAAT 1358
Db 24652 ATCTTTGAAGGCTGGATTCGTGAAGAATTTGTGTGAATTTGTTTGTGCTGGAAT 24711
Qy 1359 GTCCTTTCTTCCAACTATTGTGACAGAAAGTTTCTTAAGTCAGTAGTCTGCGCTGA 1418
Db 24712 ACTTGTGTTCTCCATCTATGTAATTAGAGAGTTGGCTGCTGTATAGTCTGGCTGG 24771
Qy 1419 CATCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGAGGCGCTCTTACATTTTGAGTT 1478
Db 24772 CATTTGTGTTCTTGTAGTCTGTATAACATCTGTCCAGGCTCTTCTGGCTTTCATAGTC 24831
Qy 1479 TCTATTGAAAGTCAGGTGAATCTTAATACATCTGCTGCTTTATATGTTAATGCTTTT 1538
Db 24832 TCTGCT--GAAAAGTCTGTGAATCTGATAGGCTTCCCTTATATGTTACTT--GACCTT 24889
Qy 1539 TTCCCTTGCATCTTTTAATATCTTCTTCTTGTCTATACCTTTAGTCAATTTGATTA 1598
Db 24890 TCTCCCTTACTGCTTTTAATATCTTCTTATTTAGTGCATTTGTTTCTGATTAATTA 24949
Qy 1599 TGCACCTGTGGGAGTTCTTTTCCGGTCCAACTATTTGTTGTTTGTATGCTTCTTTGTA 1658
Db 24950 TGTGTCGGAGAAATCTTTTCTGGTCCAGTCTATTTGAGTCTGTAGGCTTCTTTGTA 25009
Qy 1659 CCTGTAGGATCTCTTCTCAAGTTAGAAATTTCTTTTGGTTTCTTCTGAAAA 1718
Db 25010 TGATCATGGGATCTCTTTTAAAGTTGGAAAGTTTCTCTATTATTTTGTGAAGA 25069
Qy 1719 TATTTCCTGCTTTTGACCTG-----CCTTCTCCCTTCTCTCTATTCT 1764
Db 25070 TATTAGCTGGCCCTTTAAGTTGAAATCTTCAATCTCATCAATTCCTATTATCGTAGGT 25129
Qy 1765 TTGTTTTTGATAGTGTCTCGGCTTCTCGATGTTTTATGCTCGATTAATTTAGACT 1824
Db 25130 TTGTTCTTCTCAATGTGTCTGGATTAACCTGGATGTTTGAGTTAGGATCCTTTTGCAAT 25189
Qy 1825 TAAATTTTCTTGACCAAGGTATCCATTTCTTCTATCTGCTCTCACTGCTGAGATTC 1884
Db 25190 TTGTAATTTCTTGACTGTTGTGCGATGTTCTCTATGGAATCTCTGACCTGAGATTC 25249
Qy 1885 TCTCTTATCTTGTATTTCTGTCAGTGAGGCTTGTCTCTGAGGTTCCCTG--TTGGGTTT 1943
Db 25250 TCTCTTCAATTTCTTGTATTTCTGTTGCGATGCTCGCATCTATGTTCCAGATCTCTTTC 25309
Qy 1944 TTAATTTTTCATTTCCAGATTTCTTCAGTTTGGTTTGTGTTTAAATTCATTTCCA 2003
Db 25310 CTAGGAATTCATCTCCAGGCTGCTCGCTTGGGTTTCTCTTAAATGTTGCTACTTCCC 25369
Qy 2004 CTTTCAGGCTCTGAAATCTTTTACTCAATTTTCTCCAGATTT--TACATTTTCAAGG 2060
Db 25370 CTTTGTAGTATGTTTGTGTTTCAATTCATCCATCACTGTTTGGATGTTTCTGTTT 25429
Qy 2061 TTTCTTAAATGGAATTAATCAATTTCTCTTCAAGAGCTTTTATGAATTCATAAAATGTA 2120
Db 25430 TTTCTTAAATGATTTCACTGTTTGGCTGTGTTTCTGCTTTTCTTAAAGGCTCTGTA 25489

RESULT 4

ABN85752/c

ID ABN85752 standard; DNA; 173810 BP.

XX
AC
XX

DT 14-OCT-2002 (first entry)
XX Mouse chromosome 11 BAC clone RP23-270L8 SEQ ID NO 10.
DB
XX Mouse; Can 1; antiinfertility; gynaecological; infertility;
KW premature ovarian failure; menopause; Sertoli Cell only syndrome;
KW BAC clone RP23-270L8; GenBank reference AC083815; chromosome 11; ds.
XX Mus musculus.
XX US2002119929-A1.
XX 29-AUG-2002.
XX 02-NOV-2001; 2001US-00003806.
XX 03-NOV-2000; 2000US-0245872P.
XX (BISH/) BISHOP C'E.
PA (AGOU/) AGOULNIK A I.
PA (ZHUQ/) ZHU Q.
XX Bishop CB, AgoulNIK AI, Zhu Q;
XX WPI; 2002-618953/66.
XX A nucleic acid molecule (I) encoding a Can 1 polypeptide used in treating
PT infertility.
PT
XX Disclosure; Page; 45pp; English.
XX The invention relates to a nucleic acid molecule (I) encoding a Can 1
CC polypeptide. The Can 1 nucleic acid molecule is used to diagnose or treat
CC infertility or premature ovarian failure or Sertoli Cell only syndrome in
CC a mammal. The present sequence is that of a mouse Can 1 encoding
CC chromosome 11 BAC clone RP23-270L8 of the invention. Note: The present
CC sequence is not given in the printed specification but was obtained
CC through the GenBank reference AC083815
XX
SQ Sequence 173810 BP; 50474 A; 33519 C; 33994 G; 55823 T; 0 U; 0 Other;
Query Match 25.7%; Score 585.1; DB 1; Length 173810;
Best Local Similarity 62.8%; Pred. No. 0.69;
Matches 1168; Conservative 0; Mismatches 569; Indels 123; Gaps 18;
Qy 292 TCTATTTCTGATTTCTATCTTGGCTCAATTTTAACTCAGTAGTAGTTGTTGTTCC 351
Db 123634 TTTCTTTCTTATTCTCTTCTGACCAAGGATCATTGAGAAGAGTTGTTCAAGTTCC 123575
Qy 352 ATAAGTTTGAAGTTTCTGTTGTTCTTCTGTTGTTGTTGTTATCTAGATTTAAGCTGT 411
Db 123574 ATGTGAATGTGGCTTCTGTTTATTATTTGTTATTGAAGAT---CAGCCTTAGTGAT 123518
Qy 412 GGTGTCAGATAGGACATAGATTTATTTCAATTTCTTTTATCTGTCGAGACTGCTT 471
Db 123517 GGTGTCGTGATAGGATACATGGACAATTTCAATATTTTGAATCTGTTGAGCCTGAT 123458
Qy 472 TGTTTTGAATATGATTTCAATTTTGGAGA--GTTTCATAGGCTGCTGACAAGAGTAC 529
Db 123457 TGTGACCTATTATGTTGGTCAATTTTGGAGAAGTACCATGAGTGTGAGAAGAAGTAT 123398
Qy 530 AGTC--TTTGTGTTTGGTGAATAGTCTCTAAATATCTCT--AGTCCACTTGGTTTATGA 587
Db 123397 ATCCTTTGTGTTTAGGATAAATGTTTCTGTAGATATCTGTAGATCCCATTTGTTTCATCA 123338
Qy 588 CATCAGTTAGTCCACATTTCTCTGTTTCTGTTTGTGTTTGTGAGTACCTAACTGTTTGA 647
Db 123337 CTTCTGTAGTTTCAGTGTGTCCTGTTTAGTTTCT--GTTTCCAATGATCTGTCCATGTT-123279
Qy 648 GAGAATGGGATTTGAAGTAGGCCACTATCTGTGTGTAGGT--CAATATGTGATTTTAC 706
Db 123278 GAAAGTGGTGTGTGAAGTCTCCCACTATTATTGTGTAGGTGCAATGTGTGCTTTGAGC 123219

QY 707 TGTAGCTGCTGTTGTTTATGAACCTGGGTGACATGTTGTTGGTGCATAGACATTAAGA 766
Db 123218 TTTACTAAAGTTCTTTAGTGAATGGCTGCTCTGTTATTTGGAGCATAGATATTGAGA 123159
QY 767 ATTGAATGCTCTCTGTTGGTGAATTTTCCCTTGTAGTCTAGTATTTCTCCCATCT 825
Db 123158 ATTGAGATGCTCTCTGTTGGAGATTTTACCTTTGTAGTGAATGAAGTGTCCCTCTCTCT 123099
QY 826 CATCTGCTTAGTTTGGGTTTAAAGTCTAT-TAGTCAGATATTAATAAGACTGTATCGGCT 884
Db 123098 TTTTGTAGTCTTGGGTTGGAGTCAATCTATCAGATATTAGATGGCTACTCTGCT 123039
QY 885 TGTCTCTTAGGCGCATTTGCTTTAGATTA-TCTTTTCCATCTTTTACTCTTAAGTGTATGT 943
Db 123038 TGTCTCTCATACCATTTGCTTTGGAATAATGTTTTCAGCCCTTTCACTCTGAGGTAGTGT 122979
QY 944 CTATCCAT--GGTAGGTTGCTTTTGGAGTGCAGCAGTAGGATGGATCTTTGTTTTCAT 1000
Db 122978 CTATCTTTTCTCTGAGATGAGTTTCTGTAGCAGCAAAATGTTGGGTCTTTGTTGCT 122919
QY 1001 ATCCATCTGTACCCAGTATCTTTTCTAGAGAAATTAAGATCATTAGGTCATTTAGT 1060
Db 122918 AGCCAGTTTGTAGTCTATGCTTTTATTTGCGAGTTGAGA-----CCATTGATGT 122867
QY 1061 TCGAATTAATCAATGAGCAGTGTGTTGATCTTGTATCTTGCATCTGTGAGTGTGT 1120
Db 122866 TAAGAGATTAAGGAAAGTAAATGTTGCTTCTCTGTTATTTTAGTTGTTAAAGTGGCA 122807
QY 1121 GT 1180
Db 122806 TTCGTCTGT 122774
QY 1181 GT 1240
Db 122773 -----TTGAGGGAT 122765
QY 1241 TTATTAATTCATATTTTCTGTAATGGGTAAATCTTTTAGATTTGAAGTTTTCCTTAGC 1300
Db 122764 TACCTCTGTGTTTCTTAGGGCTGTGTCGGTCTTGTATTTGTTTCTGTTATT 122705
QY 1301 CTTCTTT--AGTCTGCAATTTGAAGATAGATATTTCTTACATCTGATTTTATCTTAGAAT 1358
Db 122704 ATCCCTTTGAAGGGCTGGATCTGTGAGAGATAAATCGGTGAATTTGGTTTGTCTGGAAT 122645
QY 1359 GTCTTTCTTCTCCAACTATTGTGACAGAAATTTTCTAAGTCAGTAGTCTGCGCTGA 1418
Db 122644 ACTTGGTTTCTCCCTCTATGATTAATGAGATTTGGCTGGGTATAGTACCTGGGTGC 122585
QY 1419 CATCTGTAGTCTCTTGGAGTCTGTAGCAGATCTGTGCGAGGCTTCTTACATTTTGAGTT 1478
Db 122584 AGTTTGTGTTCTTTAGTGTCTGTATAACAATCTGTCAGGCTCTTCTGCTTTCATAGTC 122525
QY 1479 TCTATTGAAAAGTCAGGTGAATTTCTAATACATCTGCTTTATATGTTAATTTGCTTTT 1538
Db 122524 TCTGTG-TGAAAATCTGTTGAATTTCTGATAGGCTGCTTTATATGTTACTT-GACCTT 122467
QY 1539 TTTCCCTTGCACTTTTAAATATCTTTCTTGTCTATATCTTATAGTATTTAGTATTTATTA 1598
Db 122466 TTTCCCTTACTGCTTTTAGTATTTCTATCTTTATTTAGTCAATTTGATTTCTGATTATTA 122407
QY 1599 TGCATGTGGGAGTTTCTTTCCGGTCCAAATCTATTTTGGTGTGTTTGTATGCTCTTGTA 1658
Db 122406 TGTGTGCGGAGGAATTTCTTTTCTGGTCCAGTCTATTTTGGAGTTCTGTAGGCTTTCTGTA 122347
QY 1659 CTTTCATAGGATCTCTTTCTCAAGGTTAGGAAATTTTCTTTTGTGTTTCTTCTGAAA 1718
Db 122346 TGTTCATATGATCTATCTTCTTAGATTTGGAGTGTTCATTAATTTTGTGGAAGA 122287
QY 1719 TATTTTCCCTGTTGTTACCTG-----CCTTCTCCCTTCTCTATTTCCT 1764
Db 122286 TGTTTGCTGGACCTTTGAGTTGAAAATCTTCAATCTCATCCACTCTTATTATCCGTACGT 122227
QY 1765 TTGGTTTTTGCATAGTGTCTGGCTCTCTGAGTGTGTTTATGCGCTGGATTTATTAGACT 1824

Db 122226 TTGGTCTTCTTATTGTTGCTCTGATTTCTCTGATATTTTGATTTAGTATCTTTTTCATT 122167
QY 1825 TAACATTTTCTTTTGACCAAGGTATCCATTTCTTCTATCTTGTCTTCACTGCTGAGATTC 1894
Db 122166 TTCCATTTTCTTTGATTTGTTGCGGATGTTCTCTATGGAATCTTCTGACCTGAGATTC 122107
QY 1895 TCTTCTATCTCTTGTATTTCTGTCAGTGGCTGTCTCTGAGTTCCTG-TTGGGTTTC 1943
Db 122106 TCTCTTCCATCTCTTGTATTTCTTGTGCTGATCTCAAACTCTATGTTCCAGATTTCTTTC 122047
QY 1944 TTAATTTTCTTCAATTTCCAGATTTCTTCAAGTGTGTTTGTGTTTATTAATTTCTATTTCCA 2003
Db 122046 CTAGGTTTCTATCTCTAGTGTGCTGCTTTGAGTTTCTTATTTGTGTTACTTCCC 121987
QY 2004 CTTTCAGGTCCTGAAATGTTTACTCATTTTCTCCC---AGTATTTACATTTTCATAGG 2060
Db 121986 TTTTTAGGTCATGATGTTTGTTCATTTCCATCCATCACCTGTTGTATGTTTTCCTCTT 121927
QY 2061 TTTCTTTATGATTTATTCATTTCTCTCTTCAAGGACCTTTTATGAATTCATAAATGTA 2120
Db 121926 TTTCTGTAAGGACTTCTACCTGTTTGATTTGTTTCTTCTTAAAGACTTGTGA 121867

RESULT 5
ABN85752
ID ABN85752 standard; DNA; 173810 BP.
XX
AC ABN85752;
XX
DT 14-OCT-2002 (first entry)
XX
DE Mouse chromosome 11 BAC clone RP23-270L8 SEQ ID NO 10.
XX
KW Mouse; Can 1; antiinfertility; gynaecological; infertility;
KW premature ovarian failure; menopause; Sertoli Cell only syndrome;
KW BAC clone RP23-270L8; GenBank reference AC083815; chromosome 11; ds.
XX
OS Mus musculus.
XX
PN US2002119929-A1.
XX
PD 29-AUG-2002.
XX
PF 02-NOV-2001; 2001US-00003806.
XX
PR 03-NOV-2000; 2000US-0245872P.
XX
PA (BISH/) BISHOP C B.
PA (AGOU/) AGOULNIK A I.
PA (ZHUQ/) ZHU Q.
XX
PI Bishop CB, AgoulNIK AI, Zhu Q;
XX
XR WPI; 2002-618953/66.
XX
XX A nucleic acid molecule (I) encoding a Can 1 polypeptide used in treating infertility.
PT
PT
PS Disclosure; Page; 45pp; English.
XX
CC The invention relates to a nucleic acid molecule (I) encoding a Can 1 polypeptide. The Can 1 nucleic acid molecule is used to diagnose or treat infertility or premature ovarian failure or Sertoli Cell only syndrome in a mammal. The present sequence is that of a mouse Can 1 encoding chromosome 11 BAC clone RP23-270L8 of the invention. Note: The present sequence is not given in the printed specification but was obtained through the GenBank reference AC083815
CC
SQ Sequence 173810 BP; 50474 A; 33519 C; 33994 G; 55823 T; 0 U; 0 Other;
Query Match 25.4%; Score 578.299; DB 1; Length 173810;
Best Local Similarity 63.2%; Pred. No. 0.74;

PF 31-JAN-2003; 2003WO-US003194.
 XX 31-JAN-2002; 2002US-00060087.
 PR 15-MAR-2002; 2002US-0364045P.
 PR 15-MAR-2002; 2002US-0364055P.
 PR 30-DEC-2002; 2002US-0436643P.
 XX
 PA (GENE-) GENE LOGIC INC.
 PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
 XX WPI; 2003-689530/65.
 DR
 XX Predicting a toxic effect of a compound, useful in identifying toxicity
 PT markers in liver tissues or cells for drug screening and toxicity assays,
 PT comprises preparing gene expression profile of tissue or cells exposed to
 PT the compound.
 PT
 PS Claim 1; SEQ ID NO 2986; 1156pp; English.
 XX
 CC The present invention relates to a method for predicting a toxic effect
 CC of a compound. The method comprises preparing a gene expression profile
 CC of a tissue or cell sample exposed to the compound, and comparing the
 CC gene expression profile to a database comprising SEQ ID 1-4925, where
 CC differential expression of the gene indicates at least one toxic effect.
 CC The method is useful for predicting at least one toxic effect of a
 CC compound, predicting hepatotoxicity or the progression of a toxic effect
 CC of a compound, identifying an agent that modulates the onset or
 CC progression of a toxic response, predicting the cellular pathways that a
 CC compound modulates in a cell, and identifying an agent that modulates at
 CC least one activity of a protein. The method and compositions of the
 CC present invention using a database of genes having liver toxin-induced
 CC differential expression, are useful in identifying toxicity markers in
 CC liver tissues or cells for drug screening and toxicity assays. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 8048 BP; 3068 A; 1764 C; 1532 G; 1682 T; 0 U; 2 Other;
 Query Match 25.4%; Score 578.2; DB 1; Length 8048;
 Best Local Similarity 62.0%; Pred. No. 2.8;
 Matches 1160; Conservative 0; Mismatches 574; Indels 136; Gaps 17;
 292 TCTATTCTTCTGATTTCTATCTTGGCTCATTTTAACTCAGTAGTGAGTGTGTTGGTTCC 351
 3329 TTCTTCTTATTTCTTCTTGGACAGGTTATCATCTGAGTAGGATTTGTTCAATTCC 3270
 352 ATAAGTTTGAAGTTTCTGTTGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 411
 3269 AGGTATATGTGGGATTC-----TTCCTTATTTGTTTAAAGACCAAGTTTGGCCCGT 3216
 412 GGTGCTCAGATAGGACATAGATATTTCAATTGCTTTTATCTGTCGAGACTTGCTT 471
 3215 GGTGCTCGATAGACGATGGGATTTCTATCTTCTGACCTGTTGAGCCCGTTT 3156
 472 TGTGTTGAATATATTTCAATTTTGGAG--AGTTTCATAGGTGCTGACAGAGAGGTAC 529
 3155 TTTGACCAATATATATGTTCAATTTTGGAGAAAGTGCCATGAGGAGCTGAGAAGAGTAT 3096
 530 AGTCTTTGTTTGGTGAATAGTCTGTA--TATCTAGGTCCACTGTTGTTATGA 587
 3095 ATCCTTTTGGTTAGGATAGAGTTCTAAATATATCCGTTAAGTCCATTTGGCTCATGA 3036
 588 CATCAGTTAGTCCAGCATTTCTCTGTTTCTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTT 647
 3035 CTTCCTTATGCTGTCACATCACTGTTT-AAATTCCTGTTTCCATGATCTGTCATTGAT 2977
 648 GAGATCGGATATGAGTAGCCACTCATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 706
 2976 GAGAGTGGTGTGTTGAAATCTCCACATTTATTTGTTGAGGTGCAATGTGTGTTTGAGC 2917
 707 TGTAGCTGTGCTGTTTATGAACTTGGGTGACATTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 766

Db 2916 TTTAGTAAGTTTCTTTTACGTATAGGTGCTTGTATTTGGGCATAGATATTTAGG 2857
 QY 767 ATTGAATGTCCTCTTGTGGGA-TTTTCTTTGATGCCCTATGTAGTATTTCCCAATCT 825
 Db 2856 AITGAGAGTTCACTTTGGTGGATTTTCTTTGATGAATGAAGTGTCTTCTTATCT 2797
 QY 826 CATCTGCTTAGTTTGGGTTTAAAGTCTA-TTAGTCAGATATTAATAAGCTGTATCGCT 884
 Db 2796 TTTTGTATGACATTTAGTTGGAAATTTGATTTTATTTAGATTTAGATGGCTACTCCAGT 2737
 QY 885 TGTCTTCTAGGCCAATTTGCTTAG-AATATCTTTTCCATCTTTTACTCTTAAGGTGATGT 943
 Db 2736 TGTCTTCTTGACCAATTTGCTTTGGAAATTTGTTTCCAGCCCTTCCACTCTGAGGTAGTGT 2677
 QY 944 CTATCCATG---GTAGGTGTCTTTTGGATGACAGATAGGATGGAATCTTGTTTTCAT 1000
 Db 2676 CTGTCTTTGTCTCTGAGGTGTGTTTCTGTAGGCAGCAAGATGCAGGCTCCTCGTTGTGT 2617
 QY 1001 ATCCATCTGTTACCCAGTATCTTTTCTAGAGAAATTAAGATCATTCAGTCAATGATGT 1060
 Db 2616 ATCCAGTTTGTAACTATGACTTTTATTTGGGAGTTGAG-----GCCATTTGATGT 2565
 QY 1061 TGAGAAATATCAATCAGCAGTGTGTTGTGATTTCTTTGTTTATCTTGCACTTTGCAAGTGTGT 1120
 Db 2564 TGAGAGATATTAAGAAATAGTGAATTTGCTTCCGTTTATATTCATATTTGAGATGAGG 2505
 QY 1121 GT 1180
 Db 2504 TTATCTTTGTGTG----- 2492
 QY 1181 GT 1240
 Db 2491 -----CTTTCATCTCTTTGTTTGTTCGAAGCAT 2459
 QY 1241 TTATTTATCATATTTTCTTGAATGTTGGTAACTTTTGTAGATTTGAAGTTTTCCTTAGC 1300
 Db 2458 TAGTTTCTTGTCTTCTTAGGTATAGCTTGCTTCTTATGTTGGCTTTTACCATTATT 2399
 QY 1301 CTTCTT--TAGTCTGCATTTGAAGATAGATATTTTACATCTCATTTTATCTTGAAT 1358
 Db 2398 ATCTTTGTAGTGTGGAATTTTGGAAAGATTTGTGTAATTTGTTTGTCTGATGGAAT 2339
 QY 1359 GTCTTTCTTCTCAACTATTGTGACAGAAAGTTTTTCTAAGTCAGTGTCTGCGCTGA 1418
 Db 2338 GTCTTGGTTTCTCCATCAATGTTAATGAGAGTTTGTGGATACAGTAACCTGGGCTGG 2279
 QY 1419 CATCTGTAGTCTTGTGAGTCTGTAGCACATCTGTGAGGCGCTTCTTACATTTGAGTT 1478
 Db 2278 CATTTGTGTTCTCTTAGGCTGTATAACATCAGTCCAGGATCTTCTGCGCTTCTATAGT 2219
 QY 1479 TCTATTGAAAAGTCAGGTGTAATTTCTAATACATCTGCTTTTATATGTTAATTTGCTTT 1538
 Db 2218 TCT-GGCGAGAGTCTGCTGTGATTTCTGATAGGTCTCCCTTTATATGTTACTT-GACCTT 2161
 QY 1539 TTTCCCTTGCATCTTTAAATATCTTTCTTTGTTCTATACATTTTGTAGTCAATTTGATTA 1598
 Db 2160 TTTCCCTTACTGTTTAAATATCTTTCTTTTATTTTGTGCGTTTGGTGTGTTGCAATTA 2101
 QY 1599 TGCAGTGTGGGAGTTTCTTTTCCGGTCCAATCTATTGTTGGTGTGTTTGTATGTTCTTTGTA 1658
 Db 2100 TGTGAGGAGAGTGTCTTTTCTGTTCCCAATCTATTGAGGTCTGTAGGCTTCTTGTG 2041
 QY 1659 CTTGATAGGCACTCTTTTCTCAAGGTTAGGAATTTTCTTTTGTGTTTCTTTGTTGAAA 1718
 Db 2040 TGTCTATGGGTATCTCTTTTGTAGGTAGGAAAGTTTCTTCTATGTTTGTGTTGAGA 1981
 QY 1719 TATTTTCCCTCTTTGACCTG-----CCTTCTCCCTCTCTCTATTCCT 1764
 Db 1980 TATTTACTTGGTCTTTGAGCTGGGAGTCTTCACTCTCTTCTATACCTATTATTCCTTAGG 1921
 QY 1765 TTGTTTTTGCATAGTGTCTCTGGCTTCTGGAATTTTATGCTCGGATTTATTTAGACT 1824

Qy	2004	CTTTCAGGTCCTGAAGAAGTTTACTCAATTTCCCTCCAGTAATTTACATTTTCATAGGTTT	2063
Db	1680	TTTTAAATTCCTTCAACTGTTT-----GATTGTGTTTTCCTGGAATTC	1638
Qy	2064	CTTTTAATGGAATTTATTCATTTTCCTCTCAAGGACCTTTTATGAATTCATAAAATCATGT	2123
Db	1637	CTTTTACGGGATTTTGTGTGTCCTCTCTATGGGCTTCTACTGTTTATTATGTTTTCCT	1578
Qy	2124	TAAGTCCTT 2133	
Db	1577	GGAATTCCTT 1568	
RESULT 8			
AAx83007/c			
ID	AAx83007 standard; DNA; 51259 BP.		
XX			
AC	AAx83007;		
XX			
DT	31-AUG-1999 (first entry)		
XX			
XX	Partial mouse WRN genomic sequence #3.		
XX	Mouse; WRN; Werner's syndrome; detection; diagnosis; autosomal;		
KW	recessive disorder; phenotype; ss.		
XX			
OS	Mus musculus.		
XX			
PN	WO9724435-A1.		
XX			
PD	10-JUL-1997.		
XX			
PF	30-DEC-1996; 96WO-US020785.		
XX			
PR	29-DEC-1995; 95US-0009409P.		
PR	29-DEC-1995; 95US-00580539.		
PR	30-JAN-1996; 96US-0010835P.		
PR	30-JAN-1996; 96US-00594242.		
PR	12-APR-1996; 96US-00632175.		
XX			
PA	(DARW-) DARWIN MOLECULAR CORP.		
XX			
PI	Oshima J, Fu Y, Yu C, Mulligan J, Schellenberg GD;		
XX			
DR	WPI; 1997-363671/33.		
XX			
PT	Isolated nucleic acid molecule encoding the WRN gene product - useful for		
PT	detection and treatment of Werner's syndrome, and related diseases.		
XX			
PS	Claim 1; Fig 7; 153pp; English.		
XX			
CC	This sequence represents a fragment of the genomic sequence containing		
CC	the coding region for the mouse WRN gene (AAx83004). The corresponding		
CC	human gene (AAx83001) encodes a protein related to Werner's syndrome. The		
CC	products can be used for the detection and treatment of Werner's syndrome		
CC	(WS), an autosomal recessive disorder with a complex phenotype, as well		
CC	as related diseases		
XX			
XX			
SQ	Sequence 51259 BP; 14533 A; 9635 C; 10266 G; 16825 T; 0 U; 0 Other;		
Query Match 24.8%; Score 565.399; DB 1; Length 51259;			
Best Local Similarity 62.1%; Pred. No. 1.4;			
Matches 1155; Conservative 0; Mismatches 576; Indels 128; Gaps 18;			
Qy	292	TCTATTCTTGATTTCTATCTTTGGCTCATTTTTAACTCAGTAGTCAGTTGTTGGTTTCC	351
Db	10412	TCTTATTTTCATTTTCTGCTTGACCAAGGTATCATTCAGTAGCGGTGTTTCAGCTTCC	10353
Qy	352	ATAGTTTCTAGATTTTCGTTTCTGTTCTGTTGTTGTTATCTAGATTTAAGCTGT	411
Db	10352	ATATGATGTCGTTCGTTTCGGTTGTTTGGTATTTAAGACCAATCTTAGCCGCTGT	10293
Qy	412	GGTGGTCAGATAGGACATAGATATTTTCAATTTGTCCTTTTCTCTCGAGCACTTGCTT	471

Db 10292 GGTCTGATGATAGGGTGCATGGGATGATTTCCATCATCTTGAATCTGTAGAGTCTGTTT 10233
Qy 472 TGTTTGAAATATGTAATTAATTTGGAGA--GTTTCATAGGCTGTGACAGAAGGTAC 529
Db 10232 TGTGACCAGCATATATGCTAGTCTTTGGAGAAGGTTCCATGAGGTGTGAGAAGGTAT 10173
Qy 530 AGTCTTTGCTTTTGGTGAATAGTCTGTAATAATCT-CTAGCTCCACTGGTTTATGAC 588
Db 10172 ATTATTTGCTTTTGGATGACATGTTCTATATAATATCTGTAGATCAATTTGGTTCAATC 10113
Qy 589 ATCAGTTAGTCCAGCATTTCTGTTTCTGTTTCTGTTTGGTGGATGACCTTAACCTGTTGGAG 648
Db 10112 ATCTGTTAGTTTCACTGTCTCTGATTTAGTCTTCTGTTTCCGTGATCTCTCCATTCGTCG 10053
Qy 649 AGAATGGGTATTTGAAGTACCCACTATCTGTGTGAGGT-CAATATGTAATTTAGCT 707
Db 10052 AGAGTGGGGTCTGAAATCTCCCACTATATTATGATCGGGTGTGATGTGTGCTTTGAGAT 9993
Qy 708 GTAGCTGTCTGTTGTTTATGAACCTGTGGGTGACATTTGTTTGGTGCATAGACATTAAGAA 767
Db 9992 TTAGTAAAGTTTCTTTTATGAATGTGGGTGCCCTTGCAATTTAGCATACATGTCAGAA 9933
Qy 768 TTGCAATGCTCTCTGTTGGAT-TTTCCTTTGATGCTATGTAGTATTTCTTCCCACTCTC 826
Db 9932 TTGAGAGTTCACTCTGGCAGATGTTCTCTTTGACCAGTATGAAGTGTCTCTTCTATCTT 9873
Qy 827 ATCTGCTTAGTTTGGGTTT-----AAGTCTATTAGTCAGATATAAATGACTGTATC 880
Db 9872 TTTTTTTTTTGTAACTTGGTTGAGAGTTGAATTTATTCATATTAGAATGCTACTCC 9813
Qy 881 GGCTGTGCTCTTAGG-----GCCATTTGCTAGAATA-TCTTTTCCATCTCTTTTACTCTA 934
Db 9812 AGCTTGTCTCTGGGAAACAACCAATTTGCTTGGAAAATTTTTCACACCTTGAACCTCTG 9753
Qy 935 AGTGATGCTCATCCAG--GTAGTGTCTTTTGGATGACAGCTAGTAGGATGATCT 991
Db 9752 AGGTAGTGTCTGCTTTGTCTAGTGTGATTTCTCTGTATGACGAAATGTGGTCC 9693
Qy 992 TGTTTTCATATCCATCTGTTTACCAGTATCTTTTCTAGAGAAATTAAGATCATTTAGT 1051
Db 9692 TGCATTACATCCAGTCTGTTAGTCTATGCTCTTCTTTTGGAAATTCG-----GTC 9642
Qy 1052 CATTTAGTTGAGAAATATCAATGACGAGTGTGTGATCTTGTA-TCTTGCACCTG 1110
Db 9641 CATTTAGTTTAAAGATATTAAGGAAAGATTTGTTACTCTCTGTTATTTTGTGTTG 9582
Qy 1111 TGAAGT 1170
Db 9581 TTAGAGTGGAAATATGTTGT 9537
Qy 1171 GT 1230
Db 9536 ----- 9537
Qy 1231 CTGGAATATTATTATTATCATATTTCTTGAATGTGGTAACTCTTTAGATTTGAAGTTT 1290
Db 9536 -TTGAAAGATTACTTTCTGTCTTCTCTAGGGTGTAGTTCCCTCCCTGTGTGTGTGTGT 9478
Qy 1291 TTTCTCTTA--GCCTTCTTTTAGTCTGCAATTTCAAGATAGATATTTCTTTTACATCTGATTT 1348
Db 9477 TCCATCTATTATCTTTTTAGAGCTGGATTTGTGGAATGATTTGTGTAATTTTGGTTT 9418
Qy 1349 ATCTTAGAATGCTTTCTTTCTCCAACTATTGTGACAGAAATTTTCTTAGTGCATAG 1408
Db 9417 GTCATGGAAATCTTGGTTTCTCCATCTATGATTAATTCAGAGTTTGTGGGTATAGTAG 9358
Qy 1409 TCTGCCCTGACATCTGTAGTCTCTTGGAGTCTGTAGACATCTGTGAGGGCTTCTTAC 1468
Db 9357 CTTGGCTGGCATTTGTGTTCTCTTTGGTCTGTATGACATCTGCCAGGATCTTCTAGC 9298
Qy 1469 ATTTTGAATTTCTATTTGGAAGTGTGAGTGTAAATTTCTAAATCATCTGCCCTTTATGTTA 1528

Db 9297 TTTCTAGTCTCTGTGTGAGAACTCTAGTCTCTATATCTGATAGGCTGTGCTTTATATGTTA 9238
Qy 1529 ATTGTCTTTTTTCCCTTGGCATCTTTTAATATTCTTTCTTTTCTTATACACTTTTAGTGAT 1588
Db 9237 CTT-GACTTCTTCCCTTACTGCTTTCAATCTTTCTTTTGTAAAGCAATTTGATGTT 9179
Qy 1589 TTGATTTATGACACTGTGGGAGTCTTTTCCGTCCAACTATTTTGGTGTGTTTGTAT 1648
Db 9178 TCAATTTATGTGACAGAGGA-ATTCTTTTCTTGTCCAGTCTATTTTGGAGTCTGGAG 9120
Qy 1649 GCTTCTTGTACTGATAGGATCTCTTCTCAAGTTTAGAAATTTTTTCTTTTGTGTT 1708
Db 9119 GTTCTTGTATGTTCTATGACATCTCTTCTTAGTTATGGAAGTTTCTTCTATAAT 9060
Qy 1709 TTTCTTGAATAATTTTCCCTGCTTTTGACCT-GCCTTCTTCCCTTCTCTATTTCC- 1763
Db 9059 TTGTTGAAGATATTTTACTGGCCCTTGGATTGGGAATCTTCTCTTCTATACCTATT 9000
Qy 1764 -----TTTGGTTTTGATAGTGTCTCTGGCTTCTCGATGTTTTATGCCCTGGATT 1814
Db 8999 ATCTTAGGTTTCATTTCTCAATGTGTCCTAATATTCTCTGATGTTTTGGTTAGGAC 8940
Qy 1815 ATTTAGACCTTAACATTTTCTTGACCAAGGTATCCATTTCTTCTATCTGTCTTCACTG 1874
Db 8939 TTTTTCATTTTGTATTTCTTGTGCTGTGTGCAATGTTTCTATGTAATCTTCTGCA 8880
Qy 1875 CTTGAGATCTCTCTTCTATCTCTTGTATCTGTGATCTGTGACGTGCTGTCTGAGTCTCT 1934
Db 8879 CTTGAGATCTCTCTTCTATCTCTTGTATCTGTTGGTGATCTGTATATATGACTGCT 8820
Qy 1935 G-TTGGTTTCTTAATTTTTTTCATTTCCAGATTCTCTCAGTTTGGTGTGTTTATTAAT 1993
Db 8819 GATCTCTTCTTAGGTTTCTGTCTCCAGCTGTGTCTCCCTTGTGATTTCTTATGTT 8760
Qy 1994 TCTATTTCCACTTTCAGTCTCGAAATGTTTACTCATTTTCCCTCC-AGTATTTACA 2050
Db 8759 TCTAGTTCCATTTTAGATCTCGATGTTTTGTTTCAITTTCTTCCCTGTTGATTTG 8700
Qy 2051 TTTTCAATAGTTTCTTATGATTTATTCATTTCTCTTCAAGGACCTTTTATGAAT 2109
Db 8699 TTTCTCTAGTCTTTAAAGATTTTGTGTTTCTTCTTTAAGGCTTCTAGTTGTTT 8641

RESULT 9

ADAO2627
ID ADA02627 standard; DNA; 47115 BP.

XX ADA02627;

XX AC

XX AC

XX 06-NOV-2003 (first entry)

XX Mouse Flt3 carcinoma associated gene, SEQ ID NO:1145.

XX Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer; breast;

XX prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;

XX gene; ds.

XX Mus sp.

XX WO2003057146-A2.

XX 17-JUL-2003.

XX 26-DEC-2002; 2002WO-US041414.

XX 26-DEC-2001; 2001US-00035832.

XX (SAGR-) SAGRES DISCOVERY.

XX Morris DW;

XX WPI; 2003-587068/55.

XX

PT New recombinant nucleic acid encoding carcinoma associated protein,
 PT useful for preparing compositions for treating carcinomas.

XX Claim 1; SEQ ID NO 1145; 245pp; English.

XX The invention relates to recombinant carcinoma associated (CA) nucleic
 CC acid sequences from mouse and human (ADA01482-ADA03094), and to
 CC recombinant carcinoma associated proteins (CAP) encoded by them. The
 CC invention also encompasses expression vectors and host cells comprising a
 CC CA nucleic acid, a polypeptide (especially an antibody) that specifically
 CC binds to the protein, and a biochip comprising CA nucleic acid or
 CC fragments thereof. The sequences of the invention were identified using
 CC oncogenic retroviruses, which insert into the genome of the host organism
 CC at random. Many of these do not carry transduced host oncogenes or
 CC pathogenic trans-acting viral genes, meaning that cancer incidence is a
 CC direct consequence of the effects of proviral integration into host
 CC protooncogenes. The CA nucleic acid sequences can be used to diagnose
 CC carcinoma (especially breast cancer, prostate cancer, lymphoma or
 CC leukaemia) or a propensity to carcinoma by determination of the sequence
 CC of a CA gene, or by determination of CA gene expression in particular
 CC tissues. CA nucleic acids, proteins and antibodies are also useful as
 CC therapeutic agents and in screening and evaluating drug candidates. The
 CC present sequence represents a specifically claimed murine CA nucleic acid
 CC sequence of the invention. Note: The complete sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 XX ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 47115 BP; 11483 A; 10181 C; 10540 G; 13012 T; 0 U; 1899 Other;

Query Match 24.4%; Score 555.699; DB 1; Length 47115;

Best Local Similarity 62.5%; Pred. No. 1.6;

Matches 1159; Conservative 0; Mismatches 563; Indels 131; Gaps 20;

292 TCTATTTCTGATTCATCTGCTGCTCATTTTAACTCAGTAGTGAGTCTTTGGTTTC 351
 30363 TTTCTTTTATTTTCTTCTTGACCAAGTATCATTTAGCAGAGCATTTGTCAGCTTCT 30422
 352 ATAAAGTTGTAAGTTTCTGTTGTTCTGTTGTTGTTGTTATCTAGATTTAAGCTGT 411
 30423 GTGTGATGTGTGCTTCTGTTGTTT---TTGTTGGTATTAAACACAGCCTTAGTCTGT 30479
 412 GGTGTCAGATAGACATAGATGATTTTCAATTTGCTTTTATCTGTCAGACTTGCCTT 471
 30480 GGTGATCTGATAGGATATAGGATTTTTCAGTCTTCTTGTATCTGTTGAGCCCTGTTT 30539
 472 TGTTTTGAATATATATTCATTTTGGAGA--GTTTCATAGGGTGTGCAAGAAGGTAC 529
 30540 TGTGACCAATTATATGTCAGTTTGAAGAAGTACTGTGAGGTCTGAGAAAGGTAT 30599
 530 AGTCTTTGTTGTTGG--TGAAATAGTCTGTAATATCT--CTAGGTCACTGTTGTTATGA 587
 30600 ATTCCTTTTGCTTTAGGATGAATGTTCTATAAATACCTGTTAGATCCATTTGCTTCATAA 30659
 588 CATCAGTTAGTCTCAGCATTTCTGTTGTTGTTGTTGTTGTTGATGACCTTAACCTGTTGA 647
 30660 TTTCGTGTTAGTTTCACTGTGCTCGATTTAG--TTTCTGTTTCCATGATCTGTCATTTGCT 30718
 648 GAGATCGGGTATTGAAGTAGCCACATCTGTTGTTGAGGT--CAATATGTGATTTTATG 706
 30719 GAGAGTGGGTGTTGAAGTCTCCCGCTATTATTGTTGTTGGGTGCAATGTTGCTTTGAGC 30778
 707 TGTAGCTGTCTGTTTATGAACTGGGTGACATTTGTTGTTGTCATAGACATTAAGA 766
 30779 TTTAGTAAAGTCTCTTTTATGACTGTGGATGCCCTTGCATTTGGAGCATAGATGTTTCA 30838
 767 ATTGCAATGTCCTCTGGTGA--TTTTCCTTTGATGCTGATGATGATTTCTCCCAATCT 825
 30839 ACTGAGAAATTCATCTTCTAGATTTTCTTTGACCAAGTATAAAGTGTCTCTCTATCT 30898
 826 CATCTGCTTAGTTTGGTTTAGTCTA--TTAGTCAGATTTAAATGACTGATCGGCT 884
 30899 TTTTGTATACTTTTGGTTGAAAGTTGATTTTATTCGATATTAGAATGATGATCTCCAGCT 30958

QY 885 TGCCTTCTTAGGCCATTTCCTTAGAATA--TCTTTTCCATCCCTTTTACTCTAAGGTGATGT 943
 DB 30959 TGTTCCTGAAAACAATTTGCTTGGAAAATTTGTTTCCAGACTTTTACTCTGAGGTCTGT 31018
 QY 944 CTATCCATG--GTAGGTTGTCTTTTGGATGACAGCAGTAGATGGATCTTGTTCAT 1000
 DB 31019 CAGTCTTGTCTAGTGGTGCATTTCTGTGTGACAGCAAAATGCTGGGTCTCTGTTATGT 31078
 QY 1001 ATCCATTTCTGTACCCAGTATCTTTTCTPAGAAATTAAGATCATTCAGTCAATGATGT 1060
 DB 31079 ATCCAGTCTGTAGTCTATGTCTTTTATTTATGGGAATTGAG-----TCCATTGATGT 31130
 QY 1061 TGAGATATCAATGACAGCTGTTTGTGGATCTTTGTTATCTTTCACATGTCAGTGTGAAGTGT 1120
 DB 31131 TAATAGATTAAGGAAAAGTGAATTTACTCTTCCTGTTATTTTGTGTTAGAGTGA 31190
 QY 1121 GT 1180
 DB 31191 TTATGTTGT 31224
 QY 1181 GT 1240
 DB 31225 -----TGAAGAT 31233
 QY 1241 TTATTATTCATATTTTCTTGAATGTGGTAAACATCTTTAGATTTGAAGTTTTCCTCCTAGC 1300
 DB 31234 ACTTCTTGTGTTTCTAGGATGTAGTTTCTTCTTCTGTAAGTGTTCATCTATAT 31293
 QY 1301 CTCTTTTAGGTCGATTTGAAGATAGATATTTCTTACATCTGATTTTATCTTGAATGT 1360
 DB 31294 CTTTGTAGGCTGATTTTATGGAAGATACCTGTGTAATTTGGTTTGTCTATGGAATAT 31353
 QY 1361 CTTTCTTCTCCAACTATTGTGACAGAAAGTTTCTTAAGTGCAGTAGTCTGGCTGACA 1420
 DB 31354 CTTGGTTTCTGTCTATGGCAATTTAGAGTTTGTCTGGGTATTAATAGTCTGGGCTGGCA 31413
 QY 1421 TCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGCAGGCGCTTCTTACATTTTCAAGTTTC 1480
 DB 31414 TTTGTGTTCTCTTAGGGTCTGTATGACATCTGCCAGGATCTTCTGGCTTTCATAGTCTC 31473
 QY 1481 TATTCGAAAGTCAGTCTTAATTAACATCTGCTTATATGTTAATTTGTTCTTTT 1540
 DB 31474 TGGT--GAGAAGTCTGTGTAATTTCTGATAGTCTGTCTTTATATGTTACTT--GACCTTTT 31531
 QY 1541 TCCCTTGATCTTTTAAATATCTTTTCTTGTCTATACCTTTTAGTGAATTCATTTATATG 1600
 DB 31532 TCCCTTACTGCTTTTAAATATTTATTTCCITTTGTAGTGCATTTGGTGTGTTGATATG 31591
 QY 1601 CACTGTGGGAGTCTTTTCCGGTCCAACTATTTTGGTGTGTTTGTATGCTTCTGTGACC 1660
 DB 31592 TGATGGGAGGAATTTCTCTCTGGTCCAGTCTATTTGGAGTCTGTAGGCTTCTGTATG 31651
 QY 1661 TTGATAGGCACTCTTCTCAAGGTGTAGGAAATTTTCTTTTGTGTTTCTTCAAAATA 1720
 DB 31652 TTATGCGCACTCTTCTTTAGGTAGGAAATTTTCTTCTATATTTGTTGTAAGATA 31711
 QY 1721 TTTTCCCTGCTTTTGAOCTGCTTCTTCCCTCCCTCC-----TCTATCTCTTT 1766
 DB 31712 TTTACTGCGCTTTAAAGTTGGGAATCTTCACTCTCTCTATAACTACTATCTCTAGGTTT 31771
 QY 1767 GGTTTTTCATAGTCTCTGCTGCTTCTGGATGTTTATGCTGCTGATTTATTTAGACTTA 1826
 DB 31772 GGTCTTCTCATTTGATCTCGAATTTCTGGAATTTTGGGTAGGAGCTTTTGTCTTTT 31831
 QY 1827 A-CATTTTCTTGACCAAGGATACCATTTCTTCTATCTTGTCTTCACTGCTGAGATTTCT 1885
 DB 31832 TGGTTTTCTTTAACTGCAAGTCCAAATGTTTCTATGATTTTCTATACCTAAGATTC-- 31890
 QY 1886 CTCCTTATCTCTGTTATCTGTGAGGCTTGTCTGAGGTTCTCTGAGGTTCTGTTGTT--CT 1944
 DB 31891 -----TCTCTGTATTTCTGTTGGTGTGATCTTGCATCTATGACTCTCTCTCTCCC 31942

QY 1945 TAAATTTTTCATCCAGAAATTCCTTCAGATTGGGTTTTGTTTATTAATTCATTTCCAC 2004
 Db 31943 TAGGTTTCTATCTCCGGTTTGTCTCCCTTTGGATTCTTTATTTGTTCTATTCCCA- 32001
 QY 2005 TTTTCAGTCCCTGAATGTTTACTCAATTTCTCCCTCC---AGTATTTACATTTTCATAGGT 2061
 Db 32002 TTTTAGATCTTTGGATAGTTTGTTCATTTCTCCGCTGTTTGAATGTTGTTTCTCTCTAA 32061
 QY 2062 TTTCTTAATGATTTATTCATTTCTCTCTCAAGACCTTTTATGAATTCATAA 2114
 Db 32062 TTTCTTAAGGAGATATTATGTCCTCTTAAAGTCCCTATCATCATCA 32114

RESULT 10
 ADB72365
 ID ADB72365 standard; DNA; 47115 BP.
 AC ADB72365;
 XX
 DE
 DT 04-DEC-2003 (first entry)
 XX
 XX Mouse Flt3 gene.
 KW mouse; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
 KW cancer; neoplasm; adenocarcinoma; sarcoma; gene.
 XX
 OS Mus sp.
 XX
 XX WO2003008583-A2.
 XX
 XX 30-JAN-2003.
 XX
 XX 26-DEC-2001; 2001WO-US051291.
 XX
 XX 02-MAR-2001; 2001US-00798586.
 PR 23-OCT-2001; 2001US-00004113.
 PR 08-NOV-2001; 2001US-00052482.
 PR 30-NOV-2001; 2001US-00997722.
 PR 20-DEC-2001; 2001US-00034650.
 XX
 XX (SAGR-) SAGRES DISCOVERY.
 PA
 PI Morris DW, Engelhard BK;
 XX
 XX WPI; 2003-239337/23.
 XX
 XX New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
 PT cancers, neoplasm, adenocarcinoma, or sarcomas.
 XX
 XX Claim 1; SEQ ID NO 193; 2304pp; English.
 XX
 CC The invention relates to a novel recombinant nucleic acid comprising a
 CC nucleotide sequence selected from any of the 660 sequences fully defined
 CC in the specification. A polynucleotide of the invention has cytostatic
 CC activity, and may have a use in gene therapy, or in a vaccine. The
 CC recombinant nucleic acids and polypeptides are useful for treating
 CC carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
 CC sarcomas. The present sequence represents a mouse gene of the invention.
 XX
 SQ Sequence 47115 BP; 11483 A; 10181 C; 10540 G; 13012 T; 0 U; 1899 Other;
 Query Match 24.4%; Score 555.699; DB 1; Length 47115;
 Best Local Similarity 62.5%; Pred. No. 1.6;
 Matches 1159; Conservative 0; Mismatches 563; Indels 131; Gaps 20;

QY 292 TCTATTTCTTGATTTCTATCTTGCTCATTTTAACTCAGTAGTAGTGTGGTTTCC 351
 Db 30363 TTTCTTTTATTTCTCTCTGACCAAGTTATCATTTGACAGACGATTTTCAGCTTCT 30422
 QY 352 ATAAAGTTGTAAGTTTCTGTTGTTTCTGTTGTTGTTGTTGTTATCTAGATTAAAGCTGT 411
 Db 30423 GTGTGATGTGTCTTCTGTTGTTT---TTGTTGTTATTTAAGACCAGCCTTAGTCTGT 30479

QY 412 GTGGTGCAGATAGGACATAGAGATTATTTCAATTTGTTCTTTTATCTGTCGAGACTTGCTT 471
 Db 30480 GGTGATCTGATAGGTTATATGGGATTATTTCACTCTTCTTTGTTATCTGTTAGGCTGTTT 30539
 QY 472 TGTGTTGAATATGTTATTCATTTTGGAGA--GTTTCATAGGTCGCTGACAGAGGTAC 529
 Db 30540 TGTGACCAATTAATGTCAGATTTTGAAGAAGTACTGTGAGGTGCTGAGAGAAGGTAT 30599
 QY 530 AGTCTTTGTTGTTTGG-TGAAATAGTCTGTAATAATCT-CTAGGTCACACTTGGTTTATCA 587
 Db 30600 ATTCCTTTGCTTTAGGATGAATGTTCTATAAATACCTGTTAGATCCATTTGGTTCATAA 30659
 QY 588 CATCAGTTAGTCCAGCATTTCTCTGTTTCGTTTTTGTGTGATGACCTAACTATGTTGA 647
 Db 30660 TTTCTGTTAGTTTCACTGTCTCCGATTAG-TTTCGTCTTCCATGATCTGCTCCATTGCT 30718
 QY 648 GAGATGGGGTATTGAAGTAGCCCATATCTCTGTGTGAGGT-CAATATGTCATTTTACG 706
 Db 30719 GAGATGGGGTGTGAAGTCTCCGCTATTATTGTGGGGTGCAATGTGTGCTTTGAGC 30779
 QY 707 TGTAGCTGTCTTGTGTTTATGAACCTTGGGTGACATTTGTTTGGTGCAATGACATTAAGA 766
 Db 30779 TTTAGTAAAGTCTCTTTTATGACTGTGGATGCCCTTGCAATTTGGAGCATAGATGTCAGA 30838
 QY 767 ATTGCAATGTCCTCTGTTGGGA-TTTTCCTTTGATGCCCTATGTAGTATTTCTCCCAATCT 825
 Db 30839 ACTGAGATTCATCTTCTTAGATTTTTCCTTTGACCAATATAAGTGTCTTCTCTATCT 30899
 QY 826 CATCTGCTTAGTTTGGGTTTAAAGTCTA-TTAGTCAGATATTAAATGACTGTATCGGCT 884
 Db 30899 TTTTGTATAACTTTTGGTTGAAAGTTGATTTTATTCGATATTAGAATGATCTCCAGCT 30958
 QY 885 TGCTTCTTAGGGCCATTTGCTTAGAATA-TCCTTTCCATCTCTTTTACTCTAAGGTGATGT 943
 Db 30959 TGTTCCTCGAAACAATTTGCTTGGAAAATTTGTTTCCAGACTTTTACTCTGAGTCTGCT 31018
 QY 944 CTATCCATG---GTAGTGTGCTTTTTCGATGTCAGCAGTGTAGTGTGATCTTTGTTTCAT 1000
 Db 31019 CAGTCTTGTCTGCTGAGGTGCAATTCCTGTGTGACAGAAATGCTGGGTCCTGTTATGT 31079
 QY 1001 ATCCATTTCTGTACCAGTATCTTTTCTAGAGAAATTAAGATCATGTAGTCAATGATGT 1060
 Db 31079 ATCCAGTCTGTAGTCTATGTCTTTTATTTGGGAATTCAG-----TCCATTGATGT 31130
 QY 1061 TGAGAAATTAATCAATGAGCAGTGTGTTGGATTCTTGTATCTTGCACTTGTGAAGTGTGT 1120
 Db 31131 TAATAGATTAAGGAAAAGTCAATTTACTTCTCTGTTTCTTTGTTGTAGAGGTGGAA 31190
 QY 1121 GT 1180
 Db 31191 TTAGT 31224
 QY 1181 GT 1240
 Db 31225 -----TGAAGATTT 31233
 QY 1241 TTAGTATTCATATTTCTTGAATGTGGGTAAACATCTTTAGATTGAAGTTTTTCTCCTAGC 1300
 Db 31234 ACTTCTGTTTCTTAGGATGTAGTTCTCTCTTGTGAAAGTGTGTTTCCATCTATTAT 31293
 QY 1301 CTTCTTTAGTCTGCAATTTGAAGATAGATTAATCTTTTACATCTGATTTTATCTTAGAATGT 1360
 Db 31294 CTTTGTAGGGCTGGATTATGGAAGATACGTGTAAATTTGGTTTGTTCATGGAATAT 31353
 QY 1361 CTTTCTTTCTCCAACTATTGTGACAGAAAGTTTTTCTAAGTCAGTGTCTGGCCTGACA 1420
 Db 31354 CTTGGTTTCTCTGTCTATGGCAATTTAGAGTTTTCTGGGTATAATAGTCTGGGCTGGCA 31413
 QY 1421 TCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGAGGGCTTTCTTACATTTTGTAGTTTC 1480
 Db 31414 TTTGTGTTCTCTTAGGGTCTGTATGACATCTGCCAGGATCTTCTGGCTTTTCATAGTCTC 31473
 QY 1481 TATTGGAAAGTCAGGTGTAAATCTAAATACATCTGCTCTTATATGTAAATGCTCTTTT 1540

Db 31474 TGGT-GAGAGTCGGTGAATTCGATAGTCTGTCTTATATGTTACTT-GACCTTTT 31531
 Qy 1541 TCCCTTGCACTTTTAATATCTTCTTCTTGTGTTCTATATCTTTAGTGAATTAATG 1600
 Db 31532 TCCCTTACTGCTTTTAATATATCTTCTTGTGTTCTATATCTTTAGTGAATTAATG 31591
 Qy 1601 CACTGTGGGAGTTCTTTTCCGGTCCAACTATTTGGTGTGTTGATGCTCTTGACC 1660
 Db 31592 TGATGGGAGGAATTTCTTCTCTGTCGAGCTATTTGAGTTCTCTAGCTCTTGATG 31651
 Qy 1661 TTGATAGGCACTCTCTTCTCAAGGTTAGGAATTTCTTTTGGTTTCTTGAATA 1720
 Db 31652 TTGATGGGCACTCTCTTCTTAGGTTAGGAAATTTCTTCTATAATTTGTTAAAGATA 31711
 Qy 1721 TTTTCCCTGCTTTGACCTGCTTCTTCCCTTCC-----TCTATTCTTT 1766
 Db 31712 TTTTCTGCCCCTTTAAGTTGGGAATCTTCACTCTCTCTATAACTATATCTTAGTTT 31771
 Qy 1767 GGTTTTTCATAGTCTCTCTGGCTTCTCTGGATGTTTATGCTCGAATTAATTTAGACTTA 1826
 Db 31772 GGTCTTCTCATGTATCTCTGGATTTCTCTGGATGTTTGGGTAGAGCTTTTGTCTTT 31831
 Qy 1827 A-CATTTCTTTGACCAAGGATCAATTTCTTATCTTGTCTTCACTGCTGAGATCT 1885
 Db 31832 TCCGTTTCTTTAACTGCACTGCCAAATGTTTCTATGTTATTTCTATACCTAAGATTC- 31890
 Qy 1886 CTCCTTCTATCTCTGATTTCTGTCAGTGAGCTGTCTCTGAGGTCTCTGTGGTT-CT 1944
 Db 31891 -----TCTCTGTATCTTGGTGATGCTTGCACTATGACTCTGATCTCTCCC 31942
 Qy 1945 TAATTTTTCATTTCCAGATTTCTCTGATTTCTGATTTGGTTTGTATTAATTTATTTCCAC 2004
 Db 31943 TAGGTTTCTATCTCCGGTTTGTCTCCCTTTGTGATTTCTTTATTTGTTCTATTTCCA- 32001
 Qy 2005 TTTGAGTCTTGAATGTTTACTCATTTTCTCTCC-----AGTATTTACATTTTCATAGT 2061
 Db 32002 TTTTGAATCTTGGATGTTTGTTCATTTCTCTGCTGTTGATTTGTTTCTCTGTAA 32061
 Qy 2062 TTTTGAATGATTTATTTCAATTTCTCTTCAAGGACCTTTTATGAATTCATA 2114
 Db 32062 TTCCTTAAGGAGATTAATTTATGCTCTTCTTAAAGTCTCTATCATCATCA 32114

RESULT 11

ID ADA02501
 XX ADA02501 standard; DNA; 96597 BP.

AC ADA02501;

XX 06-NOV-2003 (first entry)

DE Mouse Bach2 carcinoma associated gene, SEQ ID NO:1019.

KW Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer; breast;
 KW prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
 KW gene; ds.

XX Mus sp.

XX WO2003057146-A2.

XX 17-JUL-2003.

PF 26-DEC-2002; 2002WO-US041414.

PR 26-DEC-2001; 2001US-00035832.

XX (SAGR-) SAGRES DISCOVERY.

XX Morris DW;

XX WPI; 2003-587068/55.

DR

XX New recombinant nucleic acid encoding carcinoma associated protein,
 PT useful for preparing compositions for treating carcinomas.
 XX Claim 1; SEQ ID NO 1019; 245pp; English.
 XX The invention relates to recombinant carcinoma associated (CA) nucleic
 CC acid sequences from mouse and human (ADA01482-ADA03094), and to
 CC recombinant carcinoma associated proteins (CAP) encoded by them. The
 CC invention also encompasses expression vectors and host cells comprising a
 CC CA nucleic acid, a polypeptide (especially an antibody) that specifically
 CC binds to the protein, and a biochip comprising CA nucleic acid or
 CC fragments thereof. The sequences of the invention were identified using
 CC oncogenic retroviruses, which insert into the genome of the host organism
 CC at random. Many of these do not carry transduced host oncogenes or
 CC pathogenic trans-acting viral genes, meaning that cancer incidence is a
 CC direct consequence of the effects of proviral integration into host
 CC protooncogenes. The CA nucleic acid sequences can be used to diagnose
 CC carcinoma (especially breast cancer, prostate cancer, lymphoma or
 CC leukaemia) or a propensity to carcinoma by determination of the sequence
 CC of a CA gene, or by determination of CA gene expression in particular
 CC tissues. CA nucleic acids, proteins and antibodies are also useful as
 CC therapeutic agents and in screening and evaluating drug candidates. The
 CC present sequence represents a specifically claimed murine CA nucleic acid
 CC sequence of the invention. Note: the complete sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 96597 BP; 25792 A; 20376 C; 20447 G; 28650 T; 0 U; 1332 Other;

Query Match 23.1%; Score 525.898; DB 1; Length 96597;

Best Local Similarity 61.9%; Pred. No. 1.6; Mismatches 556; Indels 147; Gaps 21;
 Matches 1143; Conservative 0;

Qy 285 AAACACTCTTATTTCTGATTTCTATCTTGGTCTCAATTTTAACTCAGTAGTGAGTTGTTT 344
 Db 51129 AAAGTTTAAATTTCTTATTTCTTCAATGAGCAAGTTATCATTCAGTAGAGCGTTGTC 51188

Qy 345 GGTTCCTCAATGATTTGTAAGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTT 404

Db 51189 AGCTCCATGTATGTGTCTTCCCTGTTT---TTCTGGTATTAAGACCGGCTTT 51244

Qy 405 AAGCTGTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 464

Db 51245 ACTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 51304

Qy 465 CTTCCTTTGTTTGAATATGATTTCAATTTTGGAGA--GTTTCATAGGGTGTCTGACAAAG 522

Db 51305 CCTGTTTGTAAACCAATATATGTTGTCAGTTTGCAGAGGTACCATGAGGTCTGCGAAG 51364

Qy 523 AAGGTACAGTCTTTGTTT---GGTGAATAGTCTGTAAATA--TCTCTAGGTCCACTT 578

Db 51365 AAGGTACATTCATTTTGTCTTTAGGATGAATGTTCTATAATAATCTGCTAAACCTTTA 51424

Qy 579 GGTATATGACATCAGTTAGTCTCCAGCATTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTT 638

Db 51425 GGTCTTAACTTCTATTAGTTTCACTGTCTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTT 51483

Qy 639 ACTGTTGGAGAGATGGGGTATTGAAGTAGCCCACTATCTGTGTGAGGT-CAATATGT 697

Db 51484 TCCATTCGTGAGATGGGGTGTGAAGTCTCTACTATTTATTTGTGGGTGCAATGTGT 51543

Qy 698 GATTTTGTAGTGTAGTGTCTGTTTGTATGAATCTGGGTGACATTTGTTTGGTGCATAG 757

Db 51544 GCTTTGAGCT-----TAAGTTTCTTATGATGATGTTTCTTCTGATTTGGACACAG 51598

Qy 758 ACATTAAGAATTCGAATGCTCTTTGGTGA--TTTTCCTTTGATGCCCTATGATGATTTCT 816

Db 51599 ATCTTCAGAATTCGAGATTCATTTTGGTAGATTTTTCCTTTGACTAGTATGAAGTGTCT 51658

Qy 817 TCCCAATCTCATCTGTTAGTTTTCGGTTTAAAG-TCTATAGTCAGATTTAAATGACT 875

XX Oshima J, Fu Y, Yu C, Mulligan J, Schellenberg GD;
PI WPI; 1997-363671/33.
XX Isolated nucleic acid molecule encoding the WRN gene product - useful for
PT detection and treatment of Werner's syndrome, and related diseases.
PT Claim 1; Fig 7; 153pp; English.
XX
XX This sequence represents a fragment of the genomic sequence containing
CC the coding region for the mouse WRN gene (AA83004). The corresponding
CC human gene (AA83001) encodes a protein related to Werner's syndrome. The
CC products can be used for the detection and treatment of Werner's syndrome
CC (WS), an autosomal recessive disorder with a complex phenotype, as well
CC as related diseases
XX
SQ Sequence 51259 BP; 14533 A; 9635 C; 10266 G; 16825 T; 0 U; 0 Other;
Query Match 22.9%; Score 521.797; DB 1; Length 51259;
Best Local Similarity 60.7%; Pred. No. 2.2;
Matches 1118; Conservative 0; Mismatches 637; Indels 88; Gaps 17;
QY 232 TCTATTTCTGATTTCTATCTTGGCTCATTTTAACTCAGTAGAGTGTGTTGCTTCC 351
DB 26099 TCTTCATTTTCATTTCTGCTGACCAAGTTATCAITGAGTAGAGCGCTGTACGCTCC 26158
QY 352 ATAGTTTGTAAAGTTTCTGTTGTTTCTGTTGTTGTTGTTATCTAGATTAAAGCTGT 411
DB 26159 ATATGATGTGCTTCCGTTGTTGTTGTTGTTATTTAAACCAACCTTAGTCCGTTGT 26218
QY 412 GGTGTCAGATAGACATAGAGTATATTTCAATGTCTTTTATCTGTCGAGACTTGCTT 471
DB 26219 GGTCTGATGATAGGTCGATGGATGATTTCCATCATCTTGAATCTGTAGAAGTCTGTT 26278
QY 472 TGTITTTGAAATGATTTCAATTTGGAGA--GTTTCATAGGTCGTCACAGAGGTAC 529
DB 26279 TGTACACAGCTATATGTCAGTTTGGAGAGGTTCCATAGAGTGGCTGAGAGAGGTAT 26338
QY 530 AGTCTTTGTTTGGTGAATAGTCTGTAAATATCT--CTAGGTCACCTTGGTTTATGAC 588
DB 26339 ATTTTTCCTTTGGATCAGATGTTCTATAAATCTGTAGATCCATTTGGTTTCAATAC 26398
QY 589 ATCAGTAGTCCAGCATTTCTGTGTTTCGTTTTTGTGTAGATGACCTAACTGTTGGAG 648
DB 26399 ATCTGTTAGTTTCACTGTCTGTGCTTAGTTTCTGTTTCCGTCATCTGTGCCATTCGT 26458
QY 649 AGAATGGGATTTGAAGTAGCCCACTATCTGTGTGAGGTCAATATGTAATTTAGCTG 708
DB 26459 AGAGTGGGTCGTGAAATCTCCACTATTTATGATCAGGTATGATGTGTTGAGAT 26518
QY 709 TAGCTGTGCTGTTTATGAACCTTGGGTGACATTTGTTGGTGCATAGACATTAAGAAT 768
DB 26519 TTAGTAAAGTTTATGATGTTGGGTGCGCTTGCATTTGGAGCATAACATGTTTCAAGAT 26578
QY 769 TGCATGTCCTTCTGTCGAT--TTTCCCTTTCATGCTATGATGATTTCTTCCCAATCTCA 827
DB 26579 TGAGAGTTCACTTTGGCAGATGTTTCCCTTGGACCAATAGAGTGTCTTCTTATCTTT 26638
QY 828 TCTGCTTAGTTTGGGTTTAAAGTCT--ATTAGTCAGATATTAATAAGTACTGATCGGCTTG 886
DB 26639 TTTTGTAACTTGTGTCAGAGTTGAAATTTATTCATATTAAGATGGCTACTCCAGCTTG 26698
QY 887 CTTCCTTAGG-----GCCATTTGCTTAGAATA--TCTTTTCCATCTCTTTTACTCTCAAGTGA 940
DB 26699 TTTCTTGGAAACCAACCAATTTGCTTGGAAATTTGTTTCCAACTTGAATCTGAGGTAG 26758
QY 941 TGTCTATCCATG---GTAGGTTGTCTTTTGGATGAGCAGTAGTAGGATGGATCTGTGTTT 997
DB 26759 TGTCTGCTCTTGTGTCAGTGGGTGATTTCTGTGATGAGCAAAATGCTGGGTCTGTTTA 26818
QY 998 CATATCCATTTCTGTTACCAGATCTTTTCTTAGAGAAATTAAGATCATTTAGTGCATGA 1057
|||||

DB 26819 CACACCCAGTCTGTTAGTCTATGCTTTTTTTTGGAGAAATGA-----GTCCATTGA 26869
QY 1058 TGTGAGAAATATCAATGAGCAGTGTGTTGGAATCTTGTGATCTTGCACCTGTGGAAGTG 1117
DB 26870 TGTAAAGAGATATTAAGGAAAGTGAATGTTACTTCTCTGTTATTTT-----TGTTG 26920
QY 1118 TGTGCTG 1177
DB 26921 TTGTTAGAGGTGGAATTAATGTTTGTGGCTATCTCTTTTGGTTTGTGAAGATTCG 26980
QY 1178 TCTGTGTCGTGTGTCCTG 1237
DB 26981 TTTCTTGTCTTTTCTAGGTTAGTTTCCCTC-----CTTGTGTTGGTGT 27025
QY 1238 TATTTATTAATCATATTTCTTGAATGTGGAACATCTTAGATTGAAGATTTTCTCCT 1297
DB 27026 TTTCCATCTATTAATCTTTTGAAGCTGGAAGATTAATGTTGTAATTTGGTTTGTCA 27085
QY 1298 AGCCTTTCTTAGGTCTGCATTTTGAAGATAGATTAATCTTTTACATCTGATTTTATCT 1357
DB 27086 AAATACCTAGCAGCTTGACAGCACACCTGAACACTCTAGAACTTAAAGAAACAAATACAC 27145
QY 1358 TGTCTTTCTTCTCCAACTATTGTGACAGAAAGTTTCTTAAGTGCAGTAGTCTGCCCTG 1417
DB 27146 CCAAGAGGAGTAGACTGAGATTTGGAGTT-----TTGCTGGGCTG 27186
QY 1418 ACATCTGTAGTCTCTTTGGAGTCTGTAGCACATCTGTGACGGCCCTTCTTACATTTTGA 1477
DB 27187 GCATTTGTGTTCTTTAGGTCGTATGACATCTGCTAGGATCTTTTAGCTTTCATAGT 27246
QY 1478 TTCTATTGGAAGAGTCAGGTGTAATTTCTAAATACATCTGCTTTTATATGTTAATTTGCT 1537
DB 27247 TTTCTGTG--GAGAAGTCTGTGTAATTTCTGATAGGCTGCTTTATATGTTTACTTGA 27304
QY 1538 TTTTCCCTTGTGATCTTTTAAATATTTCTTCTTGTCTTATCTTTTATGATGATTTAT 1597
DB 27305 TTTCCATGCTGCTTTTAAATTTCTTCTTGTGTTAGTGCAATTTGCTGTTGATTT 27363
QY 1598 ATGCACTGTGGGAGTTTCTTTTCCGGTCCAATCTATTTGGTGTGTTTGTATGCTTCTGT 1657
DB 27364 ATGTGACAGAGGAAATTTCTTTTCTGTCAGTCTATTTGGAGTTCTGGAGGCTTCTGC 27423
QY 1658 ACCTTCATAGGCATCTCTTTTCAAGGTTAGGAATTTTCTTTTGGTGTGTTTCTTGA 1717
DB 27424 ATGTTATGGGCATCTGCTTTTATAGTTAGGAAGTTTCTTCTATAAATTTGTTGAAG 27483
QY 1718 ATATTTTCCCTGCTTTTGTGACCTGCTTCTTCCCTTCTCTATTC----- 1763
DB 27484 ATATTTTACGGCCCTTTGAGTTGGGAATCTTCACTCTCTTCTATACATATTAATCTTAG 27543
QY 1764 TTTGGTTTTTGCATAGTGTCTCTGGCTTCTGAGATGTTTATGCTCGGATTAATTTAGAC 1823
DB 27544 TTTGGTCTTCTCATTTGTCCTGGAATTTCTGATGTTTGGGTTAGGAGCTTTTGTGCAT 27603
QY 1824 TTAACATTTTCTTTGACCAAGGTATCCATTTCTTCTATCTTCTTCACTGCTGAGATTT 1883
DB 27604 TTTGTAATTTCTTTGACTGTGTGCAATATTTTCTATGATATCTTCTGCACTGAGATTT 27663
QY 1884 CTCTCTTCTATCTTTTGTATCTG--TCAGTGAAGGTTGTCTCTGAGGTTTCTG--TTGGGT 1941
DB 27664 CTCTCTTCTATCTTGTATCTGTTTGGTGTATGTTGCACTCTGCACTCTGATCTCTT 27723
QY 1942 TCTTAATTTTCTATTTCCAGATTTCTTCACTGTTGGGTTTGTGTTTATTAATCTTATTTTC 2001
DB 27724 TCTAGATTTTCTAACTCCAGGTTGTCTCCCTTTGTGATTTCTTTATTTGTTTCTAGTTTC 27783
QY 2002 CACTTTCCAGTCTCTGAAATGTTTTTACTCATTTTCTCTCCAGTATTTTACA---TTTTTCA 2058
DB 27784 CATTTTGTAGACTCTGATGTTTGTGTTTCTTCTTCTTCTGCTGTTTAAAGTGTGTTTCTCG 27843
QY 2059 GGTTCCTTTTAAAGGATTTATTTATTTCTTCTTCTTCAAGGACCTTT 2101
DB 27844 TAATCTGTAAGGAATTTTGTGTTTCTTCTTAAAGGGCTTCT 27886

RESULT 14
 AAD51996
 ID AAD51996 standard; DNA; 9330 BP.
 XX
 AC AAD51996;
 XX
 DT 02-MAY-2003 (first entry)
 XX
 DE Mouse Cyp3A11 gene.
 XX
 KW Transcription control element; transgenic; transgenic animal; toxicity;
 KW drug metabolism; mouse; Cyp3A11; ds.
 XX
 OS Mus sp.
 XX
 XX Location/Qualifiers
 FH Key 5104..6218
 FT misc_feature /tag= a
 FT /note= "Non repeat region"
 FT misc_feature 6792..9330
 FT /tag= b
 FT /note= "Non repeat region"
 XX
 XX WO200288305-A2.
 PN
 PD 07-NOV-2002.
 XX
 XX 11-APR-2002; 2002WO-US011770.
 PF
 XX
 XX 12-APR-2001; 2001US-0283534P.
 PR
 XX
 XX (XENO-) XENOGEN CORP.
 PA
 XX
 XX Zhang W, Contag P, Purchio A, Hashima S, Ma S, Nawotka K;
 PI
 XX
 XX WPI; 2003-120468/11.
 DR
 XX
 XX Identifying analytes that modulate expression of a reporter sequence
 PT mediated by human or mouse transcription control elements in transgenic
 PT rodent, comprises administering the analyte and monitoring expression of
 PT coding sequence.
 PT
 XX
 PS Claim 18; Fig 1B; 136pp; English.
 XX
 CC The invention relates to a method of identifying an analyte that
 CC modulates expression of a reporter sequence mediated by transcription
 CC control elements derived from a human CYP3A4 gene, in a transgenic,
 CC living rodent. The method involves administering the analyte to the
 CC rodent and monitoring expression of the sequence of interest where an
 CC effect on expression indicates that the analyte affects the expression of
 CC the reporter sequence. The method is useful for identifying an analyte
 CC that modulates expression of a reporter sequence and in isolating and
 CC identifying mouse and human transcriptional control elements associated
 CC with cytochrome expression e.g. Cyp3A11 and CYP3A4 respectively. The
 CC cells and transgenic animals containing the expression cassette may be
 CC used as models for studies involving toxicity and drug metabolism. The
 CC present sequence is mouse Cyp3A11 gene
 XX
 SQ Sequence 9330 BP; 2360 A; 1897 C; 2014 G; 3056 T; 0 U; 3 Other;
 Query Match 21.3%; Score 486.098; DB 1; Length 9330;
 Best Local Similarity 61.9%; Pred. No. 6.4;
 Matches 1022; Conservative 0; Mismatches 509; Indels 121; Gaps 17;
 416 GTCAGATAGGACATAGAGTATTATTCATTGCTTTTATCTGTCGAGACTTGCTTTGTT 475
 2503 GTCGTAGAGTGCATGGGACAAATTCATATTTTGTATCTGTTAGGCGCTGTTTGTG 2562
 476 TTGAATATGATTCAATTTGGAGA--GTTTCATAGGCTGCTGCACAGAGGTACAGTC 533
 2563 ACCAATTATATGTTAATTTTGGAGAAGGTTCCGAGGCTGCTGAGAAGTATAT---CAT 2619

QY 534 TTGTGTTTTTGGTGAATAGTCTGTAAATATCTCT-AGTCCACITGGTTTATGACATCA 592
 DB 2620 TTTGTTTTTAGGATAAAATGTTCTGTAGATACTGTCAATCCATTTGTTTATCACTTCT 2679
 QY 593 GTTAGCTCCAGCAATTTCTCTGTTTTCGTTTTTTTGTGAGATGACCTAACTGTGTGGAGAA 652
 DB 2680 GTTAGTTTCA--CTGTGCTCTGTTTAGTTTCTGTTTTTCATGATCTGTCCACTGATGAAG 2737
 QY 653 TGGGTAATTGAAGTAGCCCACTATCTGTGTGAGGT-CAATATGTAATTTTAGCTGTAG 711
 DB 2738 TGGTGTGTTGAAGTCTCCCACTATTATTGTGTGAGGTGAATGTGTGCTTTGAGCTTTAC 2797
 QY 712 CTGTGCTGTTTATGAACCTTGGGTGACATTTGTTTGTGATAGATAGACATTAAGAATTGC 771
 DB 2798 TAAAGTGCTTTTAATGAATGTCGCTGCCCTTCAATTGGAGCATAGATATCAAAATTGA 2857
 QY 772 AATGTCCTCTGGTGAATTTT-CTTTGATGCTATGATGATTTCTTCCCAATCTCATCT 830
 DB 2858 GAGTTCTCTCTGGAGCAATTTTACCTTTGATGATGATGAAGTCTCCCTCTGTTCTTTTT 2917
 QY 831 GCTTAGTTTTTGGGTTTAACTCTA-TTAGTCACATATTAAATGACTGTATCGGCTTGCTT 889
 DB 2918 GATACTTTGGTTTGAAGTTGATTTATTGATATTAGAATGGCTACCCAGCTTGTT 2977
 QY 890 CTTAGGGCCATTTGCTTAGAATATC-TTTTCCATCTTTACTCTAAAGTGATGCTATC 948
 DB 2978 CTTAGACCAATTTGCTTGGAAAATTTGTTTCCAGCCTTTCACTCTGAGTAGTGTCTGTC 3037
 QY 949 CAT--GGTAGGTGCTTTTTTGGATGCAGCAGTAGGATGGATCTGTTTTCATATCCA 1005
 DB 3038 TTTTCCCTGAGATGGTTTCTGTAAAGCAGAGAAATGTTGGTCTCTGTTGTGAGCA 3097
 QY 1006 TTCTGTTACCCAGTATCTTTTCTAGAGAAATAAGATCAATGAGTCATTGATGTTGAGA 1065
 DB 3098 GTCCTGTAGTCTATGCTTTTATTGGGAAATTGAG-----TCCATTGATATAAGA 3149
 QY 1066 ATTATCAATGAGCAGTGTGTTGGGATCTGTGTTATCTTGCACCTGTGAAGTGTGTG 1125
 DB 3150 GATATTAAAGAAAGTAATTTGTTGCTTCTATTTATTTTGTGTAGAGTTGGCAATCG 3209
 QY 1126 TGTGCTGT 1185
 DB 3210 TTTCTTTTGGCTGCTCTTTTGGCTTG----- 3237
 QY 1186 TGTGCTGT 1245
 DB 3238 -----TTGAGGAATTACTTT 3251
 QY 1246 ATTCATATTTTCTTGAATCTGGGTAAACATCTTTTAGATTGAACTTTTCTCTAGGCTTCT 1305
 DB 3252 TCTGCTTTTCTAGGGCGTGATGATCTGCTGTGATTTTTTTTCTGTATATCTT 3311
 QY 1306 TTAGGTCTGCATTTGAAGATAGATATCTTTACATCTGATTTTATCTTAGAATGCTTTC 1365
 DB 3312 GAAGGCTGGA-TTCTGGAAGATAATGTGAATTTGGTTTGTGTCATGGAATACTTTGG 3370
 QY 1366 TTTCTCAACTATTGTGACAGAAAGTTTTTCTAAGTGCAGTAGTCTGGCCTGAC--ATCT 1423
 DB 3371 TTTCTCCATCTATGTPAATTGAGAGTTTGGCGGTATGATAGCCTGGGCTGGCTTTT 3430
 QY 1424 GTAGTCTCTTGGAGTCTGTAGCACATCTGTGAGGCGCTTCTTACATTTTGTGTTTCTAT 1483
 DB 3431 TTGTTCTCTAGTGTCTGTATTAACATCTGTCCAGGCTCTTCTGGCTTTTTCATAGTCTGG 3490
 QY 1484 TGAAGAAGTCAGGTGAATCTAAATACATCTGCCCTTATGTTAATGTTGCTTTTCC 1543
 DB 3491 T-GAAAGTCTGCTGTAATCTGATAGGCTGCCCTTATATGTTACTT-GACCTTTCTCC 3548
 QY 1544 CTTGCATCTTTTAAATATCTTTTGTGTTCTATCTATCTTTTAGTCATTTGATTTATGAC 1603
 DB 3549 CGTACTGCTTTTAATATCTCTTTATTAGTGCAATTTGTTCTGATTTATTTGTGTGT 3608

Qy 1604 TGTGGGAGTTCTTTTCCGGTCCAAATCTATTGGGTTTGTATGCTTCTTTGACCTTG 1663
Db TGGGAGGAATCTCTTTTCTGGTCCAGTCTATATGGAGTTCTGTAGGCTTCTGTATGTTT 3668
Qy 1664 ATAGGCATCTCTCTCAAGGTAGGAAATTTTCTTTTGGTTTCTTTGTTTCTTGAATAATTT 1723
Db ATGGGCATGTCATCTTTTAGTTCGGAAGTTTCTTCTATAATTTTGTGGAATAATTT 3728
Qy 1724 TCCCTGCTTTTGACCTG-----CCTTCTCCCTTCCCTCTATTCTTCTTGGT 1769
Db GCTGGCCCTTTAAGTTGAAATCTTCAATCTCAATCACTCTATTATCTGTAGTTGGT 3788
Qy 1770 TTTTGCATAGTCTCTGGCTTCTGATGTTTATGCTTCACTGATTTTATAGACTTAACA 1829
Db CTCTCATGTTGCTCTGATTTCTGATGTTTATAGTAGGACCTTTTGTGTTTGTGTA 3848
Qy 1830 TTTTCTTTTGACCAAGGTATCCATTTCTTCTATCTTGTCTTCACTGCTGAGATCTCTCT 1889
Db TTATCTTTGATGTTGCTGATGTTCTCTATGGAATCTTCTGCACCTGAGATCTCTCT 3908
Qy 1890 TCTATCTCTGATCTCTGTCAGTGAGGCTTGTCTGAGGTTCTCTG-TTGGGTTCTTAAT 1948
Db TCCATCTTTTGTATCTCTGCTGATGCTCAGCTCTATGTTCCAGATTTCTTCTCTAGA 3968
Qy 1949 TTTTTCATTTCCAGATTTCTTCACTGTTTGGTTTGTATTAATCTTATTTCCACTTTC 2008
Db GTTTCATCTTCCAGGTTGCTTCACTTGGTTTCTTATTTGTTGCTACTTCCCTTTT 4028
Qy 2009 AGGTCCTGAAATGTTTACTCATTTTCTCTCC 2040
Db AGGTCAGTAGGCTTTGTTTCATTTCCATCAC 4060

RESULT 15

AAD51995
ID AAD51995 standard; DNA; 12275 BP.
XX
AC AAD51995;
XX
XX 02-MAY-2003 (first entry)
XX
DE Mouse Cyp3A11 promoter.
XX
XX Transcription control element; transgenic; transgenic animal; toxicity;
KW drug metabolism; mouse; promoter; Cyp3A11; ds.
XX
XX Mus sp.
FH Key Location/Qualifiers
FT misc_signal 1..11002
FT /tag= a
FT /note= "Transcriptional control element"
FT TATA_signal 10884..10887
FT /tag= b
FT misc_signal 10914
FT /tag= c
FT /note= "Transcription start site"
FT misc_signal 11003..11005
FT /tag= d
FT /note= "Translation initiation codon"
XX
FN WO200288305-A2.
XX
XX 07-NOV-2002.
XX
XX 11-APR-2002; 2002WO-US011770.
XX
XX 12-APR-2001; 2001US-0283534P.
XX (XENO-) XENOGEN CORP.
XX Zhang W, Contag P, Purchio A, Hashima S, Ma S, Nawotka K;
PI
XX

DR WPI; 2003-120468/11.
XX
PT Identifying analytes that modulate expression of a reporter sequence
PT mediated by human or mouse transcription control elements in transgenic
PT rodent, comprises administering the analyte and monitoring expression of
PT coding sequence.
XX
PS Claim 17; Fig 1A; 136pp; English.
XX
CC The invention relates to a method of identifying an analyte that
CC modulates expression of a reporter sequence mediated by transcription
CC control elements derived from a human Cyp3A4 gene, in a transgenic,
CC living rodent. The method involves administering the analyte to the
CC rodent and monitoring expression of the sequence of interest where an
CC effect on expression indicates that the analyte affects the expression of
CC the reporter sequence. The method is useful for identifying an analyte
CC that modulates expression of a reporter sequence and in isolating and
CC identifying mouse and human transcriptional control elements associated
CC with cytochrome expression e.g. Cyp3A11 and Cyp3A4 respectively. The
CC cells and transgenic animals containing the expression cassette may be
CC used as models for studies involving toxicity and drug metabolism. The
CC present sequence is mouse Cyp3A11 promoter
XX
SQ Sequence 12275 BP; 3054 A; 2555 C; 2636 G; 4027 T; 0 U; 3 Other;
Query Match 21.3%; Score 486.098; DB 1; Length 12275;
Best Local Similarity 61.9%; Pred. No. 5.6;
Matches 1022; Conservative 0; Mismatches 509; Indels 121; Gaps 17;
Qy 416 GTCAGATAGGACATAGAGTATTATTCAATGCTCTTTTATCTGTCGAGACTTGTCTTGT 475
Db 2503 GTCGTAGGATGTCATGGGACAATTTCAATATTTTGTATCTGTTGAGCCTGTTTGTG 2562
Qy 476 TTGAATATGATTCATTTTGGAGA-CTTTTCATAGGTCCTGCACAGAGGTACAGTC 533
Db 2563 ACAATTATATGTTAAATTTTGGAGAAGGTTCCGTGAGGTGCTGAGAATATAT---CAT 2619
Qy 534 TTTGTGTTTGGTGAATAGTCTGTAATATCTCT-AGGTCCACTTGGTTTATGACATCA 592
Db 2620 TTTGTTTAGGATAAAATGTTCTGTAGATATCTGCAATCCATTTGTTTTCATCTTCT 2679
Qy 593 GTTAGCTCAGCATTTCTCTGTTTGGTTTTTGTGAGATGACCTTAACCTGTGGAGAGAA 652
Db 2680 GTTAGTTTCA--CTGTGCTCTGTTTGTAGTTTCTGTTTTCATGATCTGTCACATGATGAAG 2737
Qy 653 TGGGGTATTGAAGTAGCCCACTATCTGTGTGAGGT-CAATATGATGATTTTAGCTGTAG 711
Db 2738 TGGTGTGTTGAAGTCTCCCACTATATTGTTGAGGTGCAATGTGCTTTGAGCTTTAC 2797
Qy 712 CTGTGCTTTGTTTATGAACCTTGGGTGACATTTGTTTGGTGCATAGACATTAAGAAATGC 771
Db 2798 TAAAGTGTCTTTAAATGAATGGCTGCCCTTGCATTTGGAGCATAGATATTCAAAATTGA 2857
Qy 772 AATGCTCTTGTGGTGAATTT-CCTTTGTAGCTGCTATGATATTCTTCCCATCTCATCT 830
Db 2858 GAGTTCCTCTTGGAGGATTTTACCTTTGATGATGATGAAGTGTCCCTCTCTTTT 2917
Qy 831 GCTTAGTTTGGGTTTAAAGTCTA-TTAGTCAAGATATTAATAATGACTGTATCGGCTTGC 889
Db 2918 GATAACTTTGTTTGGAGTTGATTTTATTGATATAGAAATGGCTTACCCAGCTTGT 2977
Qy 890 CTTAGGGCCATTTGCTTAGAATATC-TTTTCATACCTTTTACTCTAAGGTGATGCTATC 948
Db 2978 CTTACAGCAATTTGCTTGGAAAATTTGTTTCCAGGCTTTCACTCTGAGTAGTGTCTGTC 3037
Qy 949 CAT---GGTAGGTTGCTTTTGGATGCAGAGTAGGATGGATCTTGTTTTCATATCCA 1005
Db 3038 TTTTCCCTGAGATGGGTTTCTGTAGCAGAGATGTTGGGTCCTGTTTGTGTAGCCA 3097
Qy 1006 TTTCTGTACCAGTATCTTTTCTAGAGAAATTAAGATCATTTAGTGTGATGTTGAGA 1065
Db 3098 GTCTGTTAGTCTATGCTTTTATTGGGAATTTAG-----TCCATTGATTAAGA 3149

Qy	1066	ATTATCAATGAGCAGTGTGGTGTAATCTCTGTTATCTTGCACCTGTGGAAGTGTGTGTG	1128
Db	3150	GATATTAAAGAAAAGTAATTTGTTCTTCCTATTATTTTTCTGTGTAGAGTTGCCATCTCG	3209
Qy	1126	TGCTGT	1185
Db	3210	TTCTTTTGGCTGTCTTCTTTTGGCTTG-----	3237
Qy	1186	TGTGTCTTGTGTGTGTGTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1245
Db	3238	-----TTGAGGAATTAACCTT	3251
Qy	1246	ATTCAATATTTCTTGAATGTGGGTAAACAATCTTTAGATTGAAGTTTTTCTCTTAGCCTTCT	1305
Db	3252	TCCTGTCTTTCTAGGGCGTGATATCTGTCTCTATTTTTTTCTGTATTAATCTCTTT	3311
Qy	1306	TTAGGCTGTCAATTGGAAGAAGATAATCTCTTTAACAATCTGAATTTATCTAGAAATCTTTTC	1365
Db	3312	GAAGGGCTGGA-TTCTCGAAAGATAATGTGTGAATTTGGTTTTGTCTCATGGAATACTTTGG	3370
Qy	1366	TTTCTCCAACCTATTGTCAAGAAAGTTTTTCTAAGTCAGTAGTCUGGCTGAC--ATCT	1423
Db	3371	TTTTCTCAATCTATGGTAAATTTGAGAGTTTTGGCGGGTATAGTAGCCGTGGCTGGCTTTT	3430
Qy	1424	GTAGTCTTCTTGGAGTCTGTAGCACATCTGTGCAGGGCTTCTTACAAATTTGAGTTCTAT	1483
Db	3431	TTGTCTCTTCTAGTGTCTGTATAACATCTGTCCAGGCTCTCTCGGCTTTCATAGTCTCTGG	3490
Qy	1484	TGAAAAAGTCAGGTGTAATTTCTAATACATCTGCCTTTATATGTAAATGTGCTCTTTTTC	1543
Db	3491	T-GAAAAAGTCGGTGTAATTTCTGATAGGCCCTGCTTTATATGTACTT-GACCTTTCTCC	3548
Qy	1544	CTTGCACTTTTTAATPCTTTCTTGTGCTAFCTACTTTAGTGAATGATATATATGCAC	1603
Db	3549	CGTACTCTTTTTAATATCTCTCTTTATTTAGTGCAATTTGTGCTCTGATTTATGTGTGT	3608
Qy	1604	TGTGGGAGTTTCTTTTCGGTCAAATCTAATTTGGTGTGTTTTGTATGCTCTCTGTACTTG	1663
Db	3609	TGGAGGAATCTCTTTTCTGGTCCAGTCTAATATGGAGTCTCTGTAGGCTCTTGTATGTTT	3668
Qy	1664	ATAGGCATCTTTCTCAAGGTAGGAAAATTTTCTTTTTTTTGGTTTTCTTGAAAAATTTT	1723
Db	3669	ATGGCATGTCAATCTTTAGGTTGGGAAAGTTTTTCTCTATAAATTTGTTGAAAAATTTT	3728
Qy	1724	TCCCTGCTTTTGACCCTG-----CCTCTCTCCCCTCTCTCTATTCCTTTGGT	1769
Db	3729	CCTGGCCCTTAAAGTGAAAAATCTTCATCTCACTCAACTCTAATATCTGTAGTTTGGT	3788
Qy	1770	TTTTGCAATGTCTCTGGCTTCCCTGGATTTTATATGCCGTGAATATTTTGAAGTAAACA	1829
Db	3789	CTTCTCAATGTGATCCTGGATTTCCCTGGATTTTGTAGTTAGGACCTTTTGTGTTTGTGA	3848
Qy	1830	TTTTTTTGGACCAAGGTATCCATTTCTCTATCTGTGCTTTCACTGGCTGAGATTCCTCTCT	1889
Db	3849	JTATCTTTGATTGTCTGTGATGTTCTCTATATGGAATCTTCTGCACTGAGATTCCTCTCT	3908
Qy	1890	TCATCTCTTGTAATCTGTGAGTAGGCTGTCTGTAGAGTTCCGTG-TGGGTTCTTAAT	1948
Db	3909	TCCATCTTTTGTATCTCTGTTGTCTGATGCTCACGCTCATGTTTCCAGATTTCTTTCTCTAGA	3968
Qy	1949	TTTTTTCATTTCCAGATTTCTCTTCAGTTTGGTTTTGTTTTAATTAATCTATTTCCACTTTT	2008
Db	3969	GTTTCTATCTCCAGCGTTGCTCACITTTGGTTTTCTTATTTGTGTCTACTTCCCTTTTT	4028
Qy	2009	AGTCTCTGAAATGTTTTACTCATTTTCTCTCC	2040
Db	4029	AGGCTCTAGTATGGCTTTGTTCAATTTCCATCTCACTAC	4060

AC	AAP92531;
XX	
DT	16-MAY-2001 (first entry)
XX	
DE	Rat T2R04 nucleotide sequence SEQ ID NO:84.
XX	
KW	Human; rat; mouse; T2R; taste receptor; G-protein coupled receptor; KW taste transduction G-protein coupled receptor; identification; tongue; KW taste sensory neuron; taste cell; taste modulator; food; KW taste signalling pathway; ds. XX OS Rattus sp. XX FN W0200118050-A2. XX PD 15-MAR-2001. XX PF 08-SEP-2000; 2000WO-USO24821. XX PR 10-SEP-1999; 99US-00393634. PR 22-FEB-2000; 2000US-00510332. XX PA (REGC) UNIV CALIFORNIA. PA (USSH) US DEPT HEALTH & HUMAN SERVICES. XX PI Zuker CS, Adler JB, Ryba N, Mueller K, Hoon M; DR WFI; 2001-211396/21. DR P-PSDB; AAB87784. XX XX Nucleic acids encoding the T2R family of G-protein coupled taste PT receptors, useful for identifying taste modulators that can be used in PT food and pharmaceutical industries to customize taste, for e.g. to PT decrease the bitter taste of food. XX PS Claim 51; Page 195-197; 249pp; English. XX CC AAP92502 to AAP92572 represent nucleic acids which encode taste CC transduction G-protein coupled receptors designated T2R proteins. CC AAB87731 to AAB87824 represent T2R proteins, and AAB87825 to AAB87830 CC represent T2R family consensus sequences from the present invention. The CC T2R proteins are taste modulators. The nucleic acids are useful as probes CC for the identification of taste cells, as the nucleic acids are CC specifically expressed in taste cells. They also serve as tools for the CC generation of taste topographic maps that elucidate the relationship CC between the taste cells of the tongue and taste sensory neurons leading CC to taste centres in the brain. The taste modulators are useful for CC pharmacological and genetic modulation of taste signalling pathways. CC Modulatory compounds comprising T2R proteins can therefore be used in CC food and pharmaceutical industries to customise taste, for e.g. to XX decrease the bitter taste of food or drugs SQ Sequence 3449 BP; 795 A; 653 C; 702 G; 1299 T; 0 U; 0 Other;
Query Match 20.9%; Score 475.598; DB 1; Length 3449;	
Best Local Similarity 63.0%; Pred. No. 11;	
Matches 965; Conservative 0; Mismatches 459; Indels 108; Gaps 16;	
QY	589 ATCAGTTAGTCTCAGCAATTCCTGTTTCGTTCGTTTTTTTGATGACACCACTGTTGGAG 648
Db	1115 ACCTGTTAGTTTCTCGTGCTCTGTTTAG-TTTCGTTTCATGATCTGTCATTGATG 1173
QY	649 AGAATCGGTATTTGAAGTAGGCCACTATCTGTGTGTGAGGTC-ATAATGTGATTTTAGCT 707
Db	1174 AGCGTGGGCTGTGAAATCTCCGACATATGTTGTGTGTGAGATAAATGTGCTTTGAGCT 1233
QY	708 GTAGCTGTCTGTTGTTTTANGAACCTGGGTGACATTTGTTTGGTGCATAGACATTAGA 767
Db	1234 TTAGTAAGATTTCTTTTGTAATGTAGGTGCTTTTGCAITTTGGTGCATAGATATTTRAG 1293
QY	768 TTGCAATGTCCTCTGGTGGGA-TTTTCCCTTGATGCTATGTAGTATTTCCCAATCTC 826
Db	1294 TTGAGAGTTTCAGCTTGTGTGATTTTTCCTTTTGATGAATATGAAGTGTCTCTGTTATCTT 1353

QY 588 CATCAGTTAGTCCAGCAATTTCTCTGTTTCCTTTTGGTGGAGTACCTAACTGTTGGA 647
Db 14753 TTCTGTAAAGTTTCACTGTGCTCTGTTTGTAGTTTG-TGTTTCCATGATCGCCCAATTGAT 14695
QY 648 GAGATGGGTATTGAAGTACCCACATATCTGTGTGAGGT-CAATATGTGATTTTATG 706
Db 14694 GAGAGTGGGTGTTGAAGATCCCACTAGTATGTTGTAGGTGCAATGTGTGCTTGAGC 14635
QY 707 TGTAGCTGTGCTGTTTATGAACCTTGGTGACATGTTGTGTTGGTGCATAGACATTAAGA 766
Db 14634 TTTAATATGTTTCTTTTATTAATGTGGCTGCCCTTGCACTTGGATCATAGATGTTTCA 14575
QY 767 ATTGCAATGCTCTTGTGTGATTTT-----CCTTGTAGTCCCTATGATTTTCCCA 821
Db 14574 ATTGAGAGTCTTCTGTGTGATTTTGTACCCCTTTGATAAATATTCAAGTCTTTCCCT 14515
QY 822 ATCTCATCTGCTTATGTTTGGGTTTAACTCTAT-TAGTCAGATATTAAATGACTGTATC 880
Db 14514 AACITTTTGTAGTCTTTTGGTGAAGTCACTATTATTCAATATTAGAATGGTTATTCC 14455
QY 881 GCTTGTCTCTTAGGGCAATTTGCTTAGAATAT-CTTTTCCATCCTTTTACTCTAAGGTG 939
Db 14454 AGCTTGTCTTGTGAACCAATTCGTTTGGAAATATGTTTTCAGACCTTTTACTCTGAGGA 14395
QY 940 ATGCTATCCATGGTAG--GTGTCTTTTGGATGACAGTGGATGGATCTTGTCTT 996
Db 14394 GTGTTGTCTTTGACACTGAGATATGTTTCTGTATGACAGCAAAATGCTGGTCCATTT 14335
QY 997 TCATATCCATCTGTTACCCAGTATCTTTTCTAGAGAAATTAAGATCATAGTCAATG 1056
Db 14334 ATGATCCAGCTGTGTAGTCTATGCTTTTATTGGGACTTGGACCAATTGA---TCTG 14278
QY 1057 ATGTTGCAATATCAATGACAGTGTGTTGGATCTTGTATCTTGCACCTGTGAAGT 1116
Db 14277 ATGTTAAGAGATTAAGGAATAGAGATGTTGCTTCTCTGTTATTTTGTGTAGTAT 14218
QY 1117 GT 1176
Db 14217 GGACTTATGTTTGTGTGCTATATCTTTTGGTGTGT----- 14180
QY 1177 GTCTGT 1236
Db 14179 -----TGAAGAAGATTAATCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 14125
QY 1237 TTATTTATTTATTCATTTTCTTGAATGTGGTAACATCTTTAGATGAAGTTTTTCTCC 1296
Db 14124 TTTTACATCTATTAT----- 14110
QY 1297 TAGCCTTCTTTAGTCTGATTTGAAGATAGATATCTTTTACATCTGATTTTATCTTGA 1356
Db 14109 -----CCTCTGTAGGACTGCAITTTGTTTAAATCTATTGTTTAAATTTGGTTTGTGATGGA 14054
QY 1357 ATGCTCTTCTTCTCCAACTATGTGACAGAAAGTTTCTTAAAGTGCAGTAGTCTGGCT 1416
Db 14053 ATATCTGTTTCTCCATCTATGGAATATGAGATTTTCTGGGTATAGTACCTAGCT 13994
QY 1417 GACATCTGTAGTCTCTTGGAGTCTAGACATCTGTGCAGGGCTTCTTACATTTGAG 1476
Db 13993 GGCAITTTATGTTCTCTTAGGGTCTGTATGAGATCTGCCAGTATCTTCTAACTTTTCATG 13934
QY 1477 TTTCTATTGGAAGTCTAGGTGTTATCTTATACATCTGCTTATATGTTAAATTTGTTCT 1536
Db 13933 TCTCTGGT-GAGAAGTCTGGGTAATATGATAGTCTGCTTAAATTTTATTTTATTTGACCT 13875
QY 1537 TTTTTCCTTGCATCTTTTAAATTTCTTTCTTGTCTTATCTTTTACTGATTTTGAATAT 1596
Db 13874 TTTTCCCTTACTACTTTTAAATTTCTTTCTTTGTTTGTGCAATTTAGTATTTTGGATAT 13815
QY 1597 TATGCACTGTGGGAGTTCTTTTCCGGTCCAACTCA-TTTGGTGTGTTTGTATGCTTCT 1655
Db 13814 TATGTGACAAGAGGAATTTCAITTTCTGCTTCTATCTACTTTTGGAAATCTGTAGGCTTCT 13755

QY 1656 GTACCTTGATAGGCATCTCTTCTCAAGTTAGGAAATTTTCTTTTGGTTTCTTTCTTGA 1715
Db 13754 GT-----ATGGGCACTCTTCTTTTAGTTAGGAAGTTTCTTCTATATTTTATGA 13701
QY 1716 AAATATTTTCCCTGCTTTTGACCTGCCCTTCTTCCCTTCCCT----- 1756
Db 13700 AGATATTACTGGCCCTTTAACTTGGGAATCTTCACTCTCTTCTTTAGGTATAGAAGAAC 13641
QY 1757 -----CTATTCCCTTGGTGTGTTTTCATAGTGTCTCTGGC 1789
Db 13640 TATCCTTATAGGCTGTGTTCTTATATATCTTATCTTAGGCCCTTCTTATGTGCTCTGAT 13581
QY 1790 TTTCTGGATGTTTATGCTGCAATTTTATAGACTTAAATTTCTTTGACCAAGGTATC 1849
Db 13580 TTTCTGGATGTTTGGGATAGGAATTTTGTATTTGCAATTTCTTTGACTGTGTGTC 13521
QY 1850 CATTTCTTCTATCTTGTCTTCACTGCCCTGAGATTTCTTCTCTATCTCTTCTTATCTCTC 1909
Db 13520 AATGTTTCTATGGTATCTTCTGACCTGGGATCTCTCTCTATCTCTTGTATTTCTGT 13461
QY 1910 AGTGAGGCTGTCTCTGAGGTTCCCTG-TTGGGTCTTAAATTTTCAITTTCCAGATTTCC 1968
Db 13460 GGTGATGCTTATCATATGACTCTCTGATCTCTTCTTAGGTTTCTATCTCCAAGTTGG 13401
QY 1969 TTCAGTTTGGGTTTGTGTTTATTAATTTTCAATTTTCCACTTTCAGTCTCTGAAATGTTTACT 2028
Db 13400 CTCCCTTTGTGATTTCTTTTATTTGTTCTTATTTCCATCTTATAGATCTCTCAATGGTTTGT 13341
QY 2029 CATTTTCTCT---CCCAGATTTTACATTTTTCATAGGTTTCTTTAAATGGATTTAT 2078
Db 13340 CAGTCTCTTACCTGTTGTTGTTTCTTCTGTAATTTCTTTAAGGATTTT 13288

RESULT 18

ADA02777/c

ID ADA02777 standard; DNA; 96595 BP.

XX ADA02777;

XX 06-NOV-2003 (first entry)

XX Mouse Ptpkr carcinoma associated gene, SEQ ID NO:1295.

XX Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer; breast;
XX prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
XX gene; ds.

XX Mus sp.

XX WO2003057146-A2.

XX 17-JUL-2003.

XX 26-DEC-2002; 2002WO-US041414.

XX 26-DEC-2001; 2001US-00035832.

XX (SAGR-) SAGRES DISCOVERY.

XX Morris DW;

XX WPI; 2003-587068/55.

XX New recombinant nucleic acid encoding carcinoma associated protein,
XX useful for preparing compositions for treating carcinomas.

XX Claim 1; SEQ ID NO 1295; 245pp; English.

XX The invention relates to recombinant carcinoma associated (CA) nucleic
XX acid sequences from mouse and human (ADA01482-ADA03094), and to
XX recombinant carcinoma associated proteins (CAP) encoded by them. The
XX invention also encompasses expression vectors and host cells comprising a
XX CA nucleic acid, a polypeptide (especially an antibody) that specifically

binds to the protein, and a biochip comprising CA nucleic acid or fragments thereof. The sequences of the invention were identified using oncogenic retroviruses, which insert into the genome of the host organism at random. Many of these do not carry transduced host oncogenes or pathogenic trans-acting viral genes, meaning that cancer incidence is a direct consequence of the effects of proviral integration into host protooncogenes. The CA nucleic acid sequences can be used to diagnose carcinoma (especially breast cancer, prostate cancer, lymphoma or leukaemia) or a propensity to carcinoma by determination of the sequence of a CA gene, or by determination of CA gene expression in particular tissues. CA nucleic acids, proteins and antibodies are also useful as therapeutic agents and in screening and evaluating drug candidates. The present sequence represents a specifically claimed murine CA nucleic acid sequence of the invention. Note: The complete sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at <http://wipo.int/pub/published/pct> sequences.

XX	SQ	Sequence	96595 BP;	27994 A;	18068 C;	18732 G;	30038 T;	0 U;	963 Other;
		Query Match	19.9%;	Score 453.098;	DB 1;	Length 96595;			
		Best Local Similarity	60.2%;	Pred. No. 3.1;					
		Matches 1007;	Conservative	0;	Mismatches 509;	Indels 157;	Gaps 17;		
Qy	470	TTTGGTTTGGAAATATGTATTTCAATTTTGGAGA--GTTTCATAGGGTGTGACAAAGGT	527						
Db	14867	TTACTATTATTAGGTATAGTCACTTTTGAGAGAATACCATAGGTGCTCAGAAGATGGT	14808						
Qy	528	ACAGTCTTTGTGTTTGGTGGAAATAGTCTGTAAATATCTTAGGTCCTAGTGGTTTATGA	587						
Db	14807	ATATTCTTTTATTAGGATGAAGTGTATATATATAT-----ATTGGTCCATAA	14754						
Qy	588	CATCAGTTAGTCCAGCAATTTCTCTCTTTTCGTTTTTGTGTAGATGCACTAACCTGTGGA	647						
Db	14753	TTTCTGTAACTTTCACTGTGCTCTCTTTAGTTTG-TGTTTCCATGATCGCCCATTTGAT	14695						
Qy	648	GAGAAAGGGGTATTGAAGTAGCCCACTATCTGTGTGAGGT-CAATATGTCATTTTACG	706						
Db	14694	GAGAGTGGGTGTGTGAAGATCCCACTAGTATTGTGTAGGGTGCATATGTGCTCTGAGC	14635						
Qy	707	TGTAGCTGTCTTGTGTTTTATGAACTTGGGTGCACTTGTGTTGGTGCCATAGACATAAGA	766						
Db	14634	TTTAAATATGTTTCTTTTATTANTATGTGGCTGCCCTGCATTTGGATCATAGATGTCAGA	14575						
Qy	767	ATTGCAATGCTCTCTGGTGGAATTT-----CCTTTGATGCCATGTAGTAGTATCTTCCCA	821						
Db	14574	ATTGAGAGTTCTTCTTGGTAGATTGTTTACCCTTTGATAAAATATTCAACGTCTTTCCCT	14515						
Qy	822	ATCTCATCTGCTAGTTTTCGGTTTAAAGTCTAT-TAGTCAGATATTAAATGACGTATC	880						
Db	14514	AACTTTTTGAATGACATTTTGGTTGAAAGTCACATTATTCAATATTAGAAATGTTTATCC	14455						
Qy	881	GGCTTGCTCTTTAGGGCCATTTGCTTAGAATAT-CTTTTCCATCTCTTTTACTCTAAGGTG	939						
Db	14454	AGCTTGTTTCTTGAACCAATTCGCTTGGAAAATAGTTTTCCAGACCTTTTACTCTGAGTA	14395						
Qy	940	ATGTCATATCAATGGTAG--GTGTCTTTTTTGGATGACAGTAGTAGGATGGATCTTGTTT	996						
Db	14394	GTGTTTCTCTTGACACTGAGATATGTTTCCGTATATCGACAAAATGCTGGGTCCATTTT	14335						
Qy	997	TCATATCCATTCTGTTACCAGATATCTTTTCTAGAGAAATTAAGATCAATCAGTCAATG	1056						
Db	14334	ATGTATCCAGCCGTGTAGTCTATGTCTTTTATTGGGACTTGAAGCCCATGA---TGTG	14278						
Qy	1057	ATGTTGAGAATTTCAATGACAGTGTGTTGTGGATCTTGTATTCTTGCACCTTGTGAAGT	1116						
Db	14277	ATGTTAAGAGATATTAAAGGAATAGTGATCTTGCTTCCGTATTATTTTGTGTAGTGAT	14218						
Qy	1117	GT	1176						
Db	14217	GGACTTATGTTGTGTGGCTATATCTTTTGGGTTGT-----	14180						
Qy	1177	GTCTGT	1236						

XX mouse; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
 KW cancer; neoplasm; adenocarcinoma; sarcoma; gene.
 XX Mus sp.
 OS WO2003008583-A2.
 XX 30-JAN-2003.
 XX 26-DEC-2001; 2001WO-US051291.
 XX 02-MAR-2001; 2001US-00798596.
 PR 23-OCT-2001; 2001US-00004113.
 PR 08-NOV-2001; 2001US-00052482.
 PR 30-NOV-2001; 2001US-00997722.
 PR 20-DEC-2001; 2001US-00034650.
 XX (SAGR-) SAGRES DISCOVERY.
 PA Morris DW, Engelhard BK;
 XX WPI; 2003-239337/23.
 XX New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
 PT cancers, neoplasm, adenocarcinoma, or sarcomas.
 XX Claim 1; SEQ ID NO 343; 2304pp; English.
 CC The invention relates to a novel recombinant nucleic acid comprising a
 CC nucleotide sequence selected from any of the 660 sequences fully defined
 CC in the specification. A polynucleotide of the invention has cytostatic
 CC activity, and may have a use in gene therapy, or in a vaccine. The
 CC recombinant nucleic acids and polypeptides are useful for treating
 CC carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
 CC sarcomas. The present sequence represents a mouse gene of the invention.
 XX Sequence 96595 BP; 27994 A; 18068 C; 18732 G; 30838 T; 0 U; 963 Other;
 SQ

Query Match 19.9%; Score 453.098; DB 1; Length 96595;
 Best Local Similarity 60.2%; Pred. No. 3.1;
 Matches 1007; Conservative 0; Mismatches 509; Indels 157; Gaps 17;

470 TTTCTTTGAAATATGATCAATTTGGAGA--GTTTCATAGGCTCTCACAGAGGT 527
 14867 TTACTATATAGTATAGTCAGTTTGGAGAGATACCATAAGGTCTCAGAGATGGT 14808
 528 ACAGCTTTTGGTTTGGTGAATAGTCTGTAAATATCTCTAGTCCACCTTGGTTATGA 587
 14807 ATATCTTTTATTTAAGGATGAATGTTATATAATATAT--ATTTGGTCATTA 14754
 588 CATCAGTAGCTCAGCATTTCTCTGTTTGGTTTGGATGACCTAACTGTGGA 647
 14753 TTTCTGTAAGTTTCACGTGTCCTGTTTGTAGTTG-TGTTTCCATGATCGCCCATGAT 14695
 648 GAGAAATGGGTTAAGTAGTACCCACATCTGTGTGTAGGT--CAATATGTGATTTAGC 706
 14694 GAGAGTGGGTTTGAAGATCCCACTAGTATGTGTAGGTTGCAATGTGCTTTGAGC 14635
 707 TGTAGCTGTGCTGTTTATGAACCTGGGTGACATGTGTTGGTGATAGACATTAAGA 766
 14634 TTTAATATGTTTCTTTTATTAATGTGGCTGCCCTGCACTTGGATCATAGATGTTTCA 14575
 767 ATTCCAATGCTCTCTGGTGATTTT-----CTTTGATGCCATGTAGTATTTCTCCCA 821
 14574 ATTGAGAGTTCTCTTGGTAGATGTTTACCCTTTGATATAAATTCACAGCTTTCTCT 14515
 822 ATCTCATCTGCTAGTTTTCGGTTTAAGTCTAT-TAGTCAGATATTAATAATGACTGTATC 880
 14514 AACTTTTTTGTATGACTTTTGGTTGAAAGTCACATATTATTCAATATTAGAATGTTATTC 14455
 881 GGCTTGCTCTTAGGGCCATTTGCTTAGAATAT-CTTTTCCATCCCTTTTACTCTAGGTG 939

14454 AGCTTCTTTTGGAAACATTGCTTGGAAATAGTTTTCCAGACCTTTACTCTCAGGTA 14395
 940 ATGCTCTATCCATGGTAG---GTTGCTCTTTTGGATGACAGAGTAGGATGATCTGTGTT 996
 14394 GTGTTTGTCTTTTGACACTGAGATATGTTTCCCTGTATGACAGAAATGCTGGGTCCATTTT 14335
 997 TCATATCCATTTCTTTACCCAGTATCTTTTCTAGAGAAATTAAGATCAATGAGTCATTG 1056
 14334 ATGTAACAGCCCTGTAGTCTATGCTTTTATTTGGGACCTGAGCCCATTTGA---TGTG 14278
 1057 ATGTTGAGAAATATCAATGACAGAGTGTGTTGGGATCTTGTATCTTGCACCTGTGAAAT 1116
 14277 ATGTTAAGAGATATTAAAGGAATAGTGAATTTGTTGCTCTCTGTTATTTTGTGTTAGT 14218
 1117 GT 1176
 14217 GGACTTATGTTGTGTGGCTATATTTCTTTTGGGTTTGT----- 14180
 1177 GTCTGT 1236
 14179 -----TGAAGAAGATTTACTTTCTTGTCTTGTAGGGGTAGTTTGCCTTGTTTTGAA 14125
 1237 TTATTTATTTATCATATTTTCTTGAATGTGGGTAAACATCTTTAGATTGAAGTTTTCCTC 1296
 14124 TTTTACATCTATTAT----- 14110
 1297 TAGCCTTTCTTTAGGTCCTGCAATTTGAAGATAGATATTTCTTACATCTGATTTTATCTAGA 1356
 14109 ----CCTCTGTAGGACTGCAATTTGTTAAACTATTTGTTAAATTTGTTTGTTCATGGA 14054
 1357 ATGCTTTCTTTTCCCAACTATTTGACAGAAAGTTTCTTAAGTCAGTAGTCTGGCCT 1416
 14053 ATATCTTGTTTTCTCCATCTATGTAATTTAGAGTTTGTGCTGGGTATAGTAGCCTAGGCT 13994
 1417 GACATCTGTAGTCTCTCGAGTCTGTAGACATCTGTGACGGGCTCTTACATTTTGTAG 1476
 13993 GGCATTTATGTTCTTTAGGGTCTGTATGAGATCTGCCACGATCTTCTACTTCTCATAG 13934
 1477 TTTCTATTGGAAAGTCAGGTGTAATTTCTATAGATCTGCTCTTATAGTTAAATGTTGCT 1536
 13933 TCTCTGGT-GAGAAGTCTGGGTAAATATGATAGTCTGCTCTTAAATGTTATTTGACCT 13875
 1537 TTTTTCCTCTGCATCTTTTAAATATTTCTTTCTTGTCTATACTTTTGTAGTATTTGATAT 1596
 13874 TTTCCCTTACTACTTTTAAATATTTCTTTCTTGTGTCATTTAGTATTTTGTAGTAT 13815
 1597 TATGACATGTGGGAGTTTCTTTCCGGTCCAATCTA-TTTGGTGTTTTGTATGCTTCTTT 1655
 13814 TATGTGACAGAGGAATTTCAATTTCTGCTTCTATCTACTTTTGAATCTGTAGGCTTCTT 13755
 1656 GTACCTTATAGGATCTCTTCTCAGGTTAGGAATTTTCTTTTGTGTTTCTTTGTTTCTGA 1715
 13754 GT-----ATGGGCACTCTTTTCTTTAGTTAGGAAGTTTCTTCTATAAATTTTGTAGA 13701
 1716 AAATATTTTCCCTGCTTTTGTGACCTGCTTCTCCCTTCCT----- 1756
 13700 AGATATTTACTGGCCCTTTAACTTTGGGAATCTTCACTCTCTTTTAGTATAGAAGAAC 13641
 1757 -----CTATTCCTTTGGTTTGTGCATAGTGTCTCTGGC 1789
 13640 TATCCTTATGCTTGTCTTCTTATTTATCTTATCTTCTTAGGGCTTCTTATGTGCTCTGAT 13581
 1790 TTTCTGGATGTTTATGCTGCTGATTTTATAGACTTAACATTTTCTTTGACCAAGGTATC 1849
 13580 TTTCTGGATGTTTGGGATAGGAATTTTGTATTTTGGCATTTTCTTTGACTGTGTGTC 13521
 1850 CATTTCTTCTATCTGTCTTCACTGCTGAGATCTCTCTTCTTCTATCTCTTGTATCTGTCT 1909
 13520 AATGTTTCTATGATGATCTTCTGACCTGGGATTTCTCTTCTATCTCTTGTATCTTCTGT 13461
 1910 AGTAGGCTTCTCTGAGGTTTCTGTG-TTGGGTTTCTTAAATTTTCTTCTTCCAGATTTCC 1968
 13460 GGTGATGCTTACATCTATGACTCTCTGATCTCTTCTTCTAGGTTTCTATCTCCAGGTTGG 13401


```
Db 39475 -----CCTATCTCTTTTAGCTGCTGAAGATTA 39446
Qy 1243 ATTATTCATATTTCTTGAATGCGGTAAACATCTTTAGAAATGAAAGTTTCTCTAGCCCT 1302
Db 39445 CTTTCTTGCTTGTCTAGGCATAGATTCCTCTCTGTTGAAGTTTCCCTTTATTAT 39386
Qy 1303 TCTTTT--AGGTCGCATTTGAAGATAGATATCTTTACATCTGATTTTATCTTAGAATGT 1360
Db 39395 CTTTGAAGGCTCGAATTTTGAAGATATTTTGTAAATTTGTTTGTTCATGGAATAG 39326
Qy 1361 CTTTCTTCTCCAACTATTTGTGACAGAAAGTTTCTTAAAGTCAGTAGTCTGGCCTGACA 1420
Db 39325 CTTGCTTCTCCATCTATGTAATTTAGTATTTTCTGCTGGGTATAGTAACCTGAGGTGGCA 39266
Qy 1421 TCTGTAGTCTCTTGGAGTCTCTAGACATCTGTGCGAGGCGCTCTTACATTTTGAAGTTTC 1480
Db 39265 TTTGTGCTCTTATGCTTATATGACATCTGCCGAGATCTTCTAGCTTTTATAGTCTC 39206
Qy 1481 TATTGGAAAGTCAGGTGTAATTTCTATACATCTGCCCTTATATGTTAAATTTGGTCTTTT 1540
Db 39205 T-GGGGGAAGTCAGGTGTAATTTTGTATGCTTGCCTTATTTTGTACTTGAAC-TTCT 39148
Qy 1541 TCCCTTGCACTTTTAAATTTCTTCTTGTCTTATCTTTTGTCTTATCTTTTGTGATTTATG 1600
Db 39147 TCCCTTACTGCTTAGT---CGTCTTGTGTTGCAATTTGGTGTGTTTGATAAATATG 39092
Qy 1601 CACTGTGGGAGTTTCTTTCCGTCCTCAATCTATTTGGTGTGTTTGTATGCTTCTTGTACC 1660
Db 39091 TGACAGAGGAATTTCTTTTCTGTGCCAACTATTTAGAGTTCTGTAGGCTCTTGTATG 39032
Qy 1661 TTGATAGCATCTCTTCTCAAGGTTAGGAAATTTTCTTTTGTGTTTCTTGAAATA 1720
Db 39031 TTCAATGTAATCTCTTCTTTAGGTAGGGAAGTTTCTTCTATAAATTTTGTGAAGATA 38972
Qy 1721 TTTTCCCTGCTTTTGACCT-CCCTTCTCCCTTCTCTATCTCTTGTGTTTGTGATAG 1779
Db 38971 TTTACTGCCCTTTAAATTTGGCAATCTTCCCTTCTGTCTATACCTATTATCTTAGGTTT 38912
Qy 1780 TGTCT 1784
Db 38911 GGCT 38907

RESULT 22
ADC85316/c
ID ADC85316 standard; DNA; 41637 BP.
AC ADC85316;
XX
XX
XX 01-JAN-2004 (first entry)
XX
XX Human Mef2c coding sequence.
XX
XX Cytostatic; gene therapy; vaccine; cancer; carcinoma-associated gene; CA;
XX secreted; transmembrane; intracellular; ds.
XX
XX Homo sapiens.
XX
XX WO2003045230-A2.
XX
XX 05-JUN-2003.
XX
XX 02-DEC-2002; 2002WO-US038582.
XX
XX 30-NOV-2001; 2001US-00997722.
XX
XX (SAGR-) SAGRES DISCOVERY.
XX
XX PI Morris DW, Engelhard EK;
XX WPI; 2003-513603/48.
XX
```

```
PT New recombinant nucleic acid comprising a nucleotide sequence of any of
PT the carcinoma-associated (CA) genes, useful for screening for drug
XX candidates for diagnosing or treating carcinomas.
XX
XX Claim 1; SEQ ID NO 102; 983pp; English.
XX
XX The invention relates to a recombinant nucleic acid comprising a
XX nucleotide sequence selected from any of the fully defined carcinoma-
XX associated (CA) genes from the 50 tables given in the specification. The
XX CA proteins are secreted, transmembrane or intracellular proteins. The
XX recombinant nucleic acids are useful for screening for drug candidates
XX for diagnosing or treating carcinomas. Sequences given in ADC85215-
XX ADC85514 represent CA genes of the invention.
XX
XX Sequence 41637 BP; 12049 A; 8888 C; 9032 G; 11648 T; 0 U; 20 Other;
XX
XX Query Match 16.8%; Score 381.798; DB 1; Length 41637;
XX Best Local Similarity 60.7%; Pred. No. 8.7;
XX Matches 914; Conservative 0; Mismatches 467; Indels 124; Gaps 19;
XX
Qy 292 TCTATTTCTGATTTCTATCTTGGCTCATTTTAACTCAGTAGTGAG-TTGTGTTGTTTC 350
Db 40299 TTTTCTTTTATTTTCTTGGACCAATTTATCATTTAGTAGAGTGTGTTGTACCTTC 40240
Qy 351 CATAACTTTGTAAAGTTTCTGTTTCTGTTGTTGTTGTTTATCTAGATTAACTG 410
Db 40239 CACATGATGTGAGCTTTCTATTAATATGTTGTTAGAGATCAGCC---TTAGTCCT 40184
Qy 411 TGTGTGTAGATAGGACATAGAGATATTTCAATGTCTTTTATCTGTCGAGACTTGTCT 470
Db 40183 TGTGTATATAGATAGATGATGATGATTTTCAA---TCTTATATCTGTTGAGGCTGTT 40127
Qy 471 TTGTTTGTAAATATGATTTCAATTTTGGAGA--GTTTCAATAGGTCCTGCACAGAGGTA 528
Db 40126 TTGTGACCTATTATATGTCAGTTTGTGAAAAGTACCTGAGTCTGCGAAGAGATTA 40067
Qy 529 CAGTC-TTGTGTGTTGGTGAATAGTCTGTAATAATCT-CTAGGTCCACTTGGTTATG 586
Db 40066 TACCCCTTTGTTTAGGATAAAATGTTCTATAGATATCTGTTAAATCCATTTGTTTATA 40007
Qy 587 ACATCAGTTAGTCCAGCAATTTCTCTGTTCTGTTTCTGATGATGACCTAACTGTTGG 546
Db 40006 ACTTCTGTTAGTTTCACTGTCTCTGTTCAAGTTTCTCAIT-CTATGATCTGTCCATGA 39948
Qy 647 AGAGATGGGTATTGAAGTAGCCCACTATCTGTGTGTG-AGGTCAATATGTGATTTTAG 705
Db 39947 TGAGAGTGGGCTATTGAAGTCTACCACTATTATTGTGTCAGTGTCTGTGTTTGTAG 39888
Qy 706 CTGTAGCTGTCTGTTTATGAACTTGGGTGACATTTGTTTGGTGCATAGACATTAAG 765
Db 39887 CTTTAGTAAACTTTCTTAAATGATGGATTCCTTGCATTTGGAGTATAGATGTTTCA 39828
Qy 766 AATTGCAATGTCCTCTTGTGTTGATTTTCTTGTAGTCCCTATCTAGTATTCTTCCCAATCT 825
Db 39827 ATATTGAGTTTCACTTGTGATATTTTATCTT--TGAGTATGAAGTTTCCCTTCTATCT 39770
Qy 826 CATCTGCTTGTGTTTGGGTTTAAAGTCTA--TTAGTCAGATATTTAAATGACATGTAATCGGC 883
Db 39769 TTTTGTGATAACTTTAGGTTGAATTTCAATTTTATTATTATTAGATGGCTACTCCAGC 39710
Qy 884 TTGCTTCTTAGGGCCATTTTGTCTTAGAATA-TCTTTTCCATCCTTTTACTCTAAGGTGATG 942
Db 39709 TTGTTTCTTGGGACCATTTGCTTGGAAAATTTGTTTCCATCTCTTTTGTCTCTGGGTAGCT 39650
Qy 943 TCTATCCATGGTAGGTTGTCTTTTGTGATGAGCAGTAGGATGATCTTGTGTTTCAATAT 1002
Db 39649 TCTGT-CTTTGTCACTGAGGTGGTTGTATGACAGCAAAATTTTGGTCTCTGTTTATGTA 39591
Qy 1003 CCATCTGTTTACCAGTATCTTTTCTAGAGAAATTAAGATCATTTAGTCATTTGATGTTG 1062
Db 39590 CCAGTCTTTTACTCTATGTAATTTATTATTAG-----TCCATTGATATTA 39547
Qy 1063 AGAATTTATCAATGACGAGGTTTGTGGATTTCTGTTTATCTTGCACCTTGTGAAGTGTGTG 1122
```


RESULT 25
 ABX77189
 ID ABX77189 standard; DNA; 41400 BP.
 XX AC ABX77189;
 XX DT 01-MAY-2003 (first entry)
 XX DE Mouse BAC279 clone containing alpha-1-acid glycoprotein 1 gene.
 XX KW Human; ss; transgenic; drug metabolism; behaviour; gene; mouse;
 KW pharmacokinetic assay; pharmacodynamic assay; toxicology; serum albumin;
 KW alpha-acidic glycoprotein; CYP; multidrug resistance protein; MRP;
 KW uridine diphosphoglucuronosyl transferase; UGT; cytochrome P450.
 XX OS Mus sp.
 XX FN WO200283897-A1.
 XX PD 24-OCT-2002.
 XX PF 18-APR-2002; 2002WO-AU000485.
 XX PR 18-APR-2001; 2001AU-00004467.
 XX PA (GENE-) GENE STREAM PTY LTD.
 XX PI Daily JW;
 XX DR WPI; 2003-093021/08.
 XX PT New transgenic non-human animal expressing a foreign polypeptide
 PT associated with drug behavior and/or metabolism, useful for studying the
 PT behavior and/or metabolism of a drug in other animals.
 XX PS Disclosure; Page 327-349; 408pp; English.
 XX CC This invention relates to a transgenic non-human animal which may be used
 CC for assessing the behaviour and/or metabolism of a drug in another animal
 CC and which expresses a foreign polypeptide associated with drug behaviour
 CC and/or metabolism. The invention also comprises a nucleic acid construct
 CC for use in producing the above transgenic non-human animal and a method
 CC of assessing the metabolism and/or behavior of a drug in an animal of
 CC interest, comprising administering a test agent to the transgenic animal
 CC and conducting analytical tests to determine drug metabolism and/or
 CC behaviour. The transgenic animal is useful in studying drug metabolism
 CC and/or behaviour in other animals. The nucleic acid construct is useful
 CC in producing the above transgenic animal and the methods are used for
 CC producing, breeding and using transgenic animals for pharmacological
 CC (e.g. pharmacokinetic or pharmacodynamic assays) and/or toxicological
 CC studies. Nucleic acid sequences used within the invention are serum
 CC albumin; alpha-acidic glycoprotein; cytochrome P450 (CYP); uridine
 CC diphosphoglucuronosyl transferase (UGT); multidrug resistance proteins
 CC and (MRP's). The present sequence represents a DNA sequence used to
 CC create a transgenic animal within the scope of the invention
 XX SQ Sequence 41400 BP; 10267 A; 10158 C; 10127 G; 10848 T; 0 U; 0 Other;
 Query Match 13.3%; Score 302.299; DB 1; Length 41400;
 Best Local Similarity 68.0%; Pred. No. 18;
 Matches 529; Conservative 0; Mismatches 222; Indels 27; Gaps 8;
 QY 1232 TGGATTATTTATTTATTTCTTCAATTTCTTGGTAAATCTTTAGATTGAAGTTT 1291
 DB 3605 TTGAAGATTACTTTCTTCTTTTCTAGGGTGAGTTTCCCTCTCTGTGTGTTT 3664
 QY 1292 TCTCTAGCCTCTT--TAGGTCTGCATTGAAGATAGATTTCTTACATCTCATTTA 1349
 DB 3665 CCATCTATTATTTCTTGTAGGGTGGATTCTGGAAGATATTGTAGATTTCGTTT 3724
 QY 1350 TCTTAGATGTCTTTCTTCTCCACTATTGTGACAGAAAGTTTTTCTTAAGTCAGTAGT 1409

DB 3725 TCATGAATATCTTGTCTTTTCTCCATCTATGTAATTGAGAGTTTCTGCTGGGTATAGTAGC 3784
 QY 1410 CTGGCCTGACATCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGAGGGCTTCTTACA 1469
 DB 3785 CTGGCTGGCAATTTGTCTTCTTAGGGTCTGTATGACATCTGCCAGGATCTTCTGGCT 3844
 QY 1470 TTTTGAGTTTCTATTGGAAAGTCAAGTCAATCTAATATACATCTGCCCTTATATGTAA 1529
 DB 3845 TTCTAGTCTCTGGT-GAGAACTCAGGTGAATCT-CTAGTCTGCCCTTATATGTATAC 3902
 QY 1530 TTGGTCTTTTCCCTTGCATCTTTTAATATCTTTTCTTCTTGTCTATATCTTTAGTGATT 1589
 DB 3903 TT-GACCTTATCCCTTACTGCTTTTAATATCTTTTCTTGT-----TAGTGTGT-TT 3955
 QY 1590 TGATTATTATGACATCTGGGGAGTTTCTTTCCGGTCCAATCTATTGGTGTGTGTATG 1649
 DB 3956 TGAATATAATGTGATGGGAGAAATTTCTTTCTCGTCCAACTATTGGAGTCTGTAGG 4015
 QY 1650 CTTCTTGACCTTGATAGGATCTCTTCTCAAGGTAGGAAATTTTCTTTTGTGTTT 1709
 DB 4016 CTTCTTGATGTTTATGGGATCTCTTTT-AGTTAAGGACGTTTCTTCTATATTT 4074
 QY 1710 TCTTGAAATATTTTCCCTGCTTTTGACCTG-----CCTTCTTCCCTTCC 1755
 DB 4075 TGTGAAGATATTTACTGGCCCTTTAAGTTGGGATCGTCACTCTCTCTATACCTATTA 4134
 QY 1756 TCTATCTCTTGGTTTTCATAGTCTCTGGCTTCTGGATGTTTATGCTGGATTA 1815
 DB 4135 TCTTAGGTTTGGTTTCTCATGTCTCTAGATTCCCGGATGTTTGGGCTAGAGCT 4194
 QY 1816 TTTTAGACTTAACATTTTCTTGACCAAGGTATCCATTTCTTCTATCTGTCTTCACATGC 1875
 DB 4195 TTTTGTCTTTTGTGTTTCTTGTGCTCAATGTTTCTTCTATGGAATCTTCTGCAC 4254
 QY 1876 CTGAGATCTCTCTATCTCTTGTATCTGTCTGAGTGGCTGTCTCTGAGGTCTCTG 1935
 DB 4255 CTGAGATCTCTCTCTATCTCTTGTATTTTGTGTGATGCTTGCATCTATGACTCTCTG 4314
 QY 1936 -TTGGTCTCTTAATTTTCTTCCAGATTTCTTCTAGTTTGGGTTTGTATTATAA 1992
 DB 4315 ATTCTTCTAGGTTTCTATTTCCAGGTTGTCTCCCTTTGTGATTTCTTAACATA 4372
 RESULT 26
 AAS99306
 ID AAS99306 standard; DNA; 46765 BP.
 XX AC AAS99306;
 XX DT 12-MAR-2002 (first entry)
 XX DE DNA encoding aldehyde dehydrogenase 5 family, member A1 (ALDH5A1).
 XX KW Aldehyde dehydrogenase 5 family member A1; ALDH5A1;
 KW succinate-semialdehyde dehydrogenase; gene therapy; probe;
 KW antisense technology; allele specific oligonucleotide; ASO;
 KW 4-hydroxybutyric aciduria; metabolic disease; transgenic animal;
 XX KW chromosome 6p22; ds.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 FT variation /tag= a
 FT FT replace(3749,T)
 FT FT /standard name= "Single nucleotide polymorphism"
 FT FT replace(3800,G)
 FT FT /tag= b
 FT FT /standard name= "Single nucleotide polymorphism"
 FT FT replace(3829,T)
 FT FT /tag= c
 FT FT /standard name= "Single nucleotide polymorphism"
 FT FT replace(3875,G)

```

FT FT      /tag= d
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(3876,A)
FT FT      /tag= e
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(3941,T)
FT FT      /tag= f
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      4050..42765
FT FT      /tag= h
FT FT      /product= "ALDH5A1"
FT FT      /note= "Aldehyde dehydrogenase 5 family member A1.
FT FT      Specifically claimed in Claim 25"
FT FT      4050..4403
FT FT      /tag= g
FT FT      /number= 1
FT FT      replace(4059,G)
FT FT      /tag= i
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(4111,A)
FT FT      /tag= j
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(4155,C)
FT FT      /tag= k
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(4179,G)
FT FT      /tag= l
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      4404..11575
FT FT      /tag= m
FT FT      /number= 1
FT FT      replace(4430,G)
FT FT      /tag= n
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      11576..11659
FT FT      /tag= o
FT FT      /number= 2
FT FT      11660..12315
FT FT      /tag= p
FT FT      /number= 2
FT FT      12316..12486
FT FT      /tag= q
FT FT      /number= 3
FT FT      replace(12415,T)
FT FT      /tag= r
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(12422,T)
FT FT      /tag= s
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      12487..13921
FT FT      /tag= t
FT FT      /number= 3
FT FT      replace(12559,T)
FT FT      /tag= u
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(13855,G)
FT FT      /tag= v
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(13917,T)
FT FT      /tag= w
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      13922..14038
FT FT      /tag= x
FT FT      /number= 4
FT FT      replace(14021,T)
FT FT      /tag= y
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      14039..24219
FT FT      /tag= z
FT FT      /number= 4
FT FT      replace(14046,A)
FT FT      /tag= aa
FT FT      /standard_name= "Single nucleotide polymorphism"

```

```

PT variation      replace(14179,A)
PT      /tag= ab
PT      /standard_name= "Single nucleotide polymorphism"
PT exon      24220..24363
PT      /tag= ac
PT      /number= 5
PT intron      24364..29453
PT      /tag= ad
PT      /number= 5
PT intron      24598..31819
PT      /tag= ae
PT      /number= 6
PT variation      replace(29397,T)
PT      /tag= af
PT      /standard_name= "Single nucleotide polymorphism"
PT      29454..29597
PT exon      /tag= ag
PT      /number= 6
PT variation      replace(29469,A)
PT      /tag= ah
PT      /standard_name= "Single nucleotide polymorphism"
PT      31820..31978
PT exon      /tag= ai
PT      /number= 7
PT intron      31979..37049
PT      /tag= aj
PT      /number= 7
PT exon      37050..37219
PT      /tag= ak
PT      /number= 8
PT intron      37220..41171
PT      /tag= al
PT      /number= 8
PT exon      41172..41230
PT      /tag= am
PT      /number= 9
PT variation      replace(41217,C)
PT      /tag= an
PT      /standard_name= "Single nucleotide polymorphism"
PT      41231..42559
PT intron      /tag= ao
PT      /number= 9
PT variation      replace(41237,A)
PT      /tag= ap
PT      /standard_name= "Single nucleotide polymorphism"
PT      replace(41403,C)
PT      /tag= aq
PT      /standard_name= "Single nucleotide polymorphism"
PT      replace(42380,T)
PT      /tag= ar
PT      /standard_name= "Single nucleotide polymorphism"
PT      42560..42765
PT exon      /tag= as
PT      /number= 10
PT XX
PT PN WO200190119-A2.
PT XX
PT XX 29-NOV-2001.
PT PD
PT PP 21-MAY-2001; 2001WO-US016558.
PT XX
PT XX 19-MAY-2000; 2000US-0205849P.
PT PR
PT XX (GENA-) GENAISSANCE PHARM INC.
PT XX
PT XX Kliem SE, Koshy B, Tanguay DA;
PT PI
PT XX WPI; 2002-089912/12.
PT DR P-PSDB; AAU73594.
PT XX
PT XX New genetic variants of human aldehyde dehydrogenase 5 family, member A1,
PT PT ALDH5A1 gene for treating metabolic diseases and for expressing ALDH5A1
PT PT protein useful in identifying drugs to treat 4-hydroxybutyric aciduria.

```


XX Human; breast specific gene; breast cancer; differential expression;
KW cytostatic; gene therapy; gene; ss.
XX Homo sapiens.
XX WO200259271-A2.
XX 01-AUG-2002.
XX 25-JAN-2002; 2002WO-US002176.
XX 25-JAN-2001; 2001US-0263757P.
XX 23-APR-2001; 2001US-0286090P.
XX 23-MAY-2001; 2001US-0292517P.
XX (GENE-) GENE LOGIC INC.
XX Orr MS, Nation M, Diggins JC, Zeng W;
XX WPI; 2002-674803/72.
XX
XX Diagnosing breast cancer in a patient comprises detecting the level of
PT gene expression in cell or tissue samples, where a differential gene
PT expression is indicative of breast cancer.
XX
XX Claim 1; SEQ ID NO 850; 260pp + Sequence Listing; English.
XX
XX The present invention relates to methods of diagnosing breast cancer in a
CC sample, which comprise detecting the level of expression in a tissue
CC sample of two or more genes selected from those shown in ABR03867-
CC ABR1112, where a differential expression of the genes indicates breast
CC cancer. The methods are useful in diagnosing, treating, detecting the
CC progression, and in monitoring treatment of breast cancer in patients.
CC The methods are also useful as a screening tool for agents that modulate
CC the onset or progression of breast cancer. The breast cancer genes may be
CC used as diagnostic markers for the prediction or identification of the
CC malignant state of breast tissue, for confirming the type and progression
CC of cancer, and for drug screening and assays. The present sequence is a
CC coding sequence of the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
XX ftp.wipo.int/pub.published_pct_sequences
SQ Sequence 70665 BP; 22090 A; 12723 C; 13020 G; 22832 T; 0 U; 0 Other;
Query Match 12.2%; Score 277.897; DB 1; Length 70665;
Best Local Similarity 52.5%; Pred. No. 17;
Matches 895; Conservative 0; Mismatches 726; Indels 83; Gaps 14;
322 TTTAACTCAGTGTGAGTGTGTTGGTTCCATAAGTTTGTAAAGTTTCTGTGTTTCTGT 381
49114 TTTCAATCAGAGCAGGTATTTAATTTCCATGTATTGCAATGGTTTGGAGTTCC--- 49058
382 TGTGTGTTGTTATCTAGATTTAAGCTGTGGTGTGATAGACATAGACATAGATTTATTT 441
49057 TTTAGGAGTTGATTTCCAGTTTATTCACATGTGCTGTGAGAGAGTGCCTGATGTAATTT 48998
442 CAATGTCTTTATCTGTGCGAGCTGTGTTTGAATATGTATTCAATTTGGAGA 501
48997 CAATTTCTTAAATTTAAGGTCATTTATGGCCCATCACAATAGTCTATCTTGAGA 48938
502 GTTTCATAGGTTGTGACAGAAGGTACAGTCT-TTGTGTTTGGTGAATAGTCTGTAA 560
48937 AAGTTCATGTGCTATGAATCATGTGTATCTGCAGTTTTTGGATGAATGTTATGTAT 48878
561 ATAATCT-----CTAGGTCCACTTGGTTTATGACATCAGTTAGTCCAGCAT 607
48877 ATATATATATATATATCTGTGTAAGTCCATTTGTTCCAAAGTTATAGTTTAAATCCATTGTT 48818
608 TCTCTGTTTCGTTTTTCTGAGATGACCTAACTGTTGGAGAGAAATGGGTATTGAAGTA 667
48817 TCTTTGTTGCCCTTTCTT-TCTTGTATGACCTCTCTAAATGCTCTCAGTGGAGTTTGAAGTC 48759

QY 668 GCCCACTATCTGTGTG-TGAGGTCAATATGTGATTTTAGCTGTAGCTGTGCTGTTTAT 726
Db 48758 CCCCCTATTATTGTGTGCTGTATCTCAATTTCTTAAGTCTATTAGTAATGTTTAA 48699
QY 727 GAACTTGGGTGACATTGTGTTTGGTGCATAGACATTAAGAAATGCAATCTCTCTGGT 786
Db 48698 ACATTGAGATCTCCAGTGTAGATATATGTTTAGATTCGATATTTTCCCTGTG 48639
QY 787 GATTTTCCCTTTGATGCTATGTAGTATTTCTCCCAATCTCATCTGCTTAGTTTGGGTT 846
Db 48638 GACGGGCAATTTGCCAATATATAATGTCCATCTTTGTCTCTTTAAATATTGTTGCTTA 48579
QY 847 AAGTCTATTAG-TCAGATATTAAATGACTGTATCGGCTTGTCTTCTTAGGGCCATTGCT 905
Db 48578 AAGTTTGTGTTTATCTGATATAAGAAATAGCTCCCGCTCACTTTTGTGTGCCATTGCA 48519
QY 906 TAGAATATCTTTTCCATCCTTTTACTCTAAAGGTGATGCTATCCATGGTAGTGTCTTT 965
Db 48518 TGAATGCTTTTCTACCTCTTAAGTTTATGTGAGTCTTATGTGTTA-CATAAGTCT 48460
QY 966 TTGGATGAGCAGTAGGATGATCTGTTTTCATATCCATTCTGTATCCCAAGTATCTTT 1025
Db 48459 CCTGAAGGCAGCAGATATTTGGTTTCATGAGTTCTTATCCATTCTGCACTTCTGTGCTTT 48400
QY 1026 TTCTAGAGAAATTAAGATCAATTGAGTCATTGATGTTGAGAAATATCAATGACAGTGT 1085
Db 48399 TAAAGTGAGCATTTAG-----GCCATTACATTCAATGATGTTTGAATGTAGG 48348
QY 1086 GTGGATTTCTTTATCTTCACCTGTGAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1145
Db 48347 TACCAATGGTTAAATCATGCTCT-----TTGTGCTGTGTGTGTGTGTGTGTGTGT 48295
QY 1146 TGT 1205
Db 48294 TTTTTTTTCTTTTAACTGATTTTGTGTTTATAGTCTGTGTGTGTGTGTGTGTGTGT 48235
QY 1206 TCTCTCCCTCTTTGATTTTGGCTGGAATATTTATTTATCATATATTTCTTGAATGT 1265
Db 48234 AAGAGTCTGTGTTGATGGGTTTCCAGATTTGTTTCAAGAT----- 48192
QY 1266 GGTAAACATCTTTAGATGAAAGTTTTCCTAGGCTTCTTTAGGTCTCCATTGAAAT 1325
Db 48191 -----TTAGAGCTCTTTTTCACAATTTCTGTAGTGTGTGTGTGTGTGTGTGT 48148
QY 1326 AGATATTTTACATCTGATTTTATCTTAGAATGCTTTTCTTCTCCAACTATTGTGACA 1385
Db 48147 GCAATTCACAGATCAATTTGTTGCTGAAATGACTGTATCTTCTCATGTATGATG 48088
QY 1386 GAAAGTTTTTCTAAGTGCAGTGTGTGGCTGCACATCTGTAGTCTCTTGGAGTCTGTAGC 1445
Db 48087 CTTAGATTCACGTGATACAAAATTTCTGGCTGATAATGT- -TTGCTTTGAGCGGTTGA 48030
QY 1446 ACATCTGTGCGAGGCTTCTTACATTTTCAATTTTCAATTTTCTTCCCTTGCATCTTTTAAATATTTCT 1505
Db 48029 AGATAGGCGCAATCCCTTCTAGCTGTAGAGTTTCTGCTGAGATATCTGCTGTTAATCT 47970
QY 1506 AATACATCTGCTCTTATATGTTAATGGTCTTTTTCCTTGCATCTTTTAAATATTTCTTT 1565
Db 47969 AATAGTTTACCCTTTATAGTTTACCTGGT-GCTTCTGTCTCACAGCTCTTAAGATCTTT 47911
QY 1566 CTTTGTCTATACTTTTGTAGTGTGATTTATGCACTGTGGGAGTGTCTTTTCCGGT 1625
Db 47910 CTTTGTCTTAACTTTGGATAAAGTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 47851
QY 1626 CCAATCTATTGTGTGTTTGTATGCTTTCTGTACCTTTGATAGGCTCTCTTTCTCAAGT 1685
Db 47850 TGAATTTCCAGGTGTTCTTTGTGCTTCTGTATTTGCAATGCTAGGTCTCTAGCAAGC 47791
QY 1686 TAGGAAATTTTCTTTTGTGTTTCTTGAATAATTTTCCCTGCTTTT-GACCTGCTTT 1744
Db 47790 CAGGCAAGTTTCTTCTGATTTTCCCAATATGTTTCCCAAGCTTTTGAATTTGCTTA 47731

Qy	1745	CTTCCC-----TTCTCTATTCTCTTTGGTTTTTGCATGTGCTCTGGCTTCCT	1794
Db	47730	CTTCTTCAGGTACATGTTTATTCTTAGGTTCATTTAACACAATCCCAGACTTCCT	47671
Qy	1795	GGATGTTTTATGCTCGGATTATTATTAGACTTAACATTTTCTTTGACCAAGGTATCCATTT	1854
Db	47670	GGAGGCTTAGTTCATATTTCTATTCTTTTTCTTTCTTTGTGATTCAGTTAATT	47611
Qy	1855	CTTCTATCTGTGTTTCACCTGCCTGCAGATTCCTCTCTCTATCTCTTGTATTTCTGTCAGTGA	1914
Db	47610	CAAAGACCTTGCTTCAAGTCTGAAATTCCTTCTTCTATTATTCAATTCCTATTGCTGA	47551
Qy	1915	GGCTTCTCTCAGGTTCTGTGTGGTTCCTTAATTTTTTTCATTTCCAGATTTCCCTCAGT	1974
Db	47550	GACITTCACAGACATTTTCAATTTCTAAAGATGTCGCAGTTTCCCAAAATTTTTTATT	47491
Qy	1975	TTGGGTTTTGTTTATTAAATTCAT	1998
Db	47490	GTTTTTCTTTAAACTACCTGTTT	47467

RESULT 30	
AAA35003	
ID	AAA35003 standard; DNA; 56583 BP.
XX	
XX	AAA35003;
XX	
DT	28-JUL-2000 (first entry)
XX	
DE	Human adenosine receptor related polynucleotide SEQ ID NO:2692.
XX	
KW	Human; adenosine receptor; low adenosine antisense oligonucleotide;
KW	phosphorothioate; impaired respiration; inflammation; allergy;
KW	allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
KW	antiallergic; anticholinergic; cytotatic; analgesic; impaired airway;
KW	lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
KW	respiratory distress syndrome; pain; cystic fibrosis; emphysema;
KW	pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
KW	cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200009525-A2.
XX	
PD	24-FEB-2000.
XX	
PF	03-AUG-1999; 99WO-US017712.
XX	
PR	03-AUG-1998; 98US-0095212P.
XX	
PA	(UYEC-) UNIV EAST CAROLINA.
XX	
PI	Nyce JW;
XX	
WPI	2000-205971/18.
XX	
XX	New antisense oligonucleotides useful for treating e.g. pulmonary
PT	vasoconstriction, inflammation, allergies, asthma, hypertension,
PT	bronchitis, emphysema, respiratory distress syndrome, ischemia or
PT	cancers.
XX	
XX	Disclosure; Page 903-916; 1343pp; English.
PS	
CC	The present invention describes a new composition comprising an antisense
CC	nucleotide (ON) with low adenosine (up to 15%), which targets
CC	nucleic acids involved in bronchoconstriction, allergies, and/or
CC	inflammation. The ON can have antiinflammatory, antiallergic,
CC	anticholinergic, cytotatic and analgesic activities. The compositions are
CC	useful for the treatment of diseases associated with inflammation,
CC	impaired airways, including lung disease and diseases whose secondary
CC	effects afflict the lungs of a subject. They can be used for treating
CC	e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma,
CC	impaired respiration, respiratory distress syndrome, pain, cystic

QY 1068 TATCAATGAGCAGTGTGTTGGAATCTTGTATATGCACTGTGGAAGTGTGTGTGTG 1127
DB 30834 TAATATTGAATGAGGACTGTGTTTCAATCATGCTCT----- 30874
QY 1128 TG 1187
DB 30875 -----TGT 30914
QY 1188 TGTCTGT 1247
DB 30915 TAACTTATATTTTGTTCATAGTCTCTATGATTTATGCTTTAAGAAGTCTGTGTTT 30974
QY 1248 TCATATTTTCTGTAATGTGGTAACATCTTTAGATTTGAAGTGTGTTTCTCCAGCTGCTTT 1307
DB 30975 TATGTGTGTTCCAGATTTGTTTCA-----AGATTTAGAGATCTCTTTTAGCAGTCTTGT 31028
QY 1308 AGGCTGCTGATTTGAAGATAGATATCTTACATCTGATTTTACTGTAGATGCTTCTT 1367
DB 31029 AATGGTGTGTTTCTTAATGGGAAATCTCTCAGCAATTTGTTGTGAAATGACTGATC 31088
QY 1368 TCTCAACTATGTGACAGAAAGTTTTTCTAAGTGCAGTAGTCTGGCCTGACATCTGTAG 1427
DB 31089 TTTCTCTTCAATATGATACCTAGTTTCTCTGGATACAAATCTCTGGCTGATTAATTT 31148
QY 1428 TCTCTGGAGTCTGTAGCAGATCTGTGAGGGCTTCTTACATTTTGAGTTTCTATTTGA 1487
DB 31149 TGTGTGAGGAGGCTGAAGTAGGTCCTCCGCTCCCTCTCTGCTGTAGGGTTTCTGTGTAG 31208
QY 1488 AAGCTCAGGTGTAATCTTAATACATCTGCTCTTATATGTTAAATGGTCTTTTCCCTTG 1547
DB 31209 AAA-TCTGCTGCTAATCTGATAGGCTTCTTTATAGTAACTGGT-GCTTCTGCTCA 31266
QY 1548 CATCTTTTAAATCTTCTTGTGTTCTAATCTTTTAGTGAATTTGATTAATGACATGTG 1607
DB 31267 CAGCTCTTAAGATCTTCTTGTGCTTAAGTTTGGATAACCTGATGACAAAGTGCTAG 31326
QY 1608 GGGAGTTTCTTTCCGGTCAACTATTTTGGTGTGTTGTATGCTCTGTACCTTGTATAG 1667
DB 31327 GTGAATATCTTTTGGATGAATTTCCAGGTGTTCTTGTGCTTCTGTATTTGGACAT 31386
QY 1668 GCACTCTCTTCTCAAGTTAGGAATTTTCTTTTGGTTTCTTGGAAATATTTTCCC 1727
DB 31387 CTAGGTCTCTGCAAGCCAGGAAGGTTTCTCGATTAATCCCTTAAACATGTTTCCA 31446
QY 1728 TGCTTTTGACCTGCTTCTTCCCTTCTCTATCC-----TTTGGTTTGTGA 1776
DB 31447 GGCTTTTGAATCTCTCTTCTTCAGGTACACTGATTAATCTTAAGTTTGGTCAATTAA 31506
QY 1777 TAGTGTCTCTGGCTTCTGATGTTTATGCTGGAATTTTAGACTTAACATTTTCTT 1836
DB 31507 CATAAATCCAGATTTCTGGAGGCTTGTTCATATTTCTTATCTTTTCTTGTCTT 31566
QY 1837 TGACCAAGGTATCCATTTCTTCTATCTGTTCTTCACTGCTGAGATCTCTCTTATCT 1896
DB 31567 TGCTGGATGGGTAATTCAT-----AGACCTTAAGCTCTGAATTTCTTATTTACTT 31620
QY 1897 CTGTATTTCTCAGTGAGGCTGCTCTGAGGTTCCTGT-----TGGGTCTTAA 1947
DB 31621 GTTCAGTTCTATGCTGAGACTTCCAGAAATTCACATTTCTGAAGTGTGTCNAAG 31680
QY 1948 TTTTTCATTTCCAGATTTCTTCCAGTTGGGTTTGTGTTTATTAATTTCTATTTCCACTTT 2007
DB 31681 TTTCTCAGTTTGTGACTGTGTTTCTTTAAGCTATCTATTTCCATGAATTTTCTCCCT 31740
QY 2008 CAGGCTCTGAAATGTTTACTCATTTTCCCT 2037
DB 31741 TCACTTCTGTATCATTTTGTGAATCTTCCCT 31770

RESULT 31

AAF21125

ID AAF21125 standard; DNA; 56583 BP.

XX AAF21125;
XX AC
XX DT
XX DT
DE 14-MAR-2001 (first entry)
XX Human low adenosine antisense oligonucleotide related sequence #2692.
XX Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
KW human; airway disorder; bronchoconstriction; lung inflammation;
KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;
KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cytosolic;
KW respiratory obstruction; pulmonary obstruction; impeded respiration;
KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;
KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
KW cancer; ss.
XX Homo sapiens.
XX WO200062736-A2.
XX 26-OCT-2000.
XX 24-MAR-2000; 2000WO-US008020.
XX 06-APR-1999; 99US-0127958P.
XX (UYEC-) UNIV EAST CAROLINA.
XX (NYCE/) NYCE J W.
XX Nyce JW;
XX WPI; 2000-679539/66.
XX Low adenosine (A) content antisense oligonucleotides which do not trigger
PT adenosine receptors during metabolism, useful e.g. for treating cancers
PT and respiratory obstructions.
XX Disclosure; Page 978-991; 1592pp; English.
XX The present invention describes low adenosine (A) content antisense
CC oligonucleotides and compositions (I) comprising them. In the antisense
CC oligonucleotides the A is replaced by a 'Universal' or alternative base.
CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.
CC The antisense oligonucleotides and (I) can be used to down-regulate the
CC expression and or activity of target polypeptides associated with
CC lung/respiratory disorders and malignancies, such as stimulating and
CC activating peptide factors and transmitters, transcription factors,
CC immunoglobulins and antibodies, antibody receptors, cytokines and
CC chemokines, endogenously produced specific and non-specific enzymes,
CC binding proteins, adhesion molecules and their receptors, cytokine and
CC chemokine receptors, adenosine receptors, bradykinin receptors, central
CC nervous system (CNS) and peripheral nervous and non-nervous system
CC receptors, CNS and peripheral nervous and non-nervous system peptide
CC transmitters, defensins, growth factors, vasoactive peptides and
CC receptors, binding proteins and malignancy associated proteins. The
CC antisense oligonucleotides may be used in this way to treat disorders
CC including respiratory obstruction (especially pulmonary obstruction
CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or
CC surfactant hypoproduction which are associated with a disease or
CC condition selected from pulmonary vasoconstriction, inflammation,
CC allergies, asthma, impeded respiration, respiratory distress syndrome
CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
CC fragments and antisense oligonucleotides used in the exemplification of
CC the present invention
XX Sequence 56583 BP; 16610 A; 11169 C; 11927 G; 16877 T; 0 U; 0 Other;

PD	31-OCT-2002.	
XX		
XX	23-APR-2002; 2002WO-US013135.	
XX		
XX	24-APR-2001; 2001US-0286137P.	
XX		
XX	(EPIG-) EPIGENESIS PHARM INC.	
XX		
PI	Nyce JW, Li Y, Sandrasagra A, Katz E, Pabalan J, Aguilar D;	
PI	Miller S, Tang L, Shahabuddin S;	
XX		
XX	WPI; 2003-229219/22.	
XX		
PT	Pharmaceutical composition for treating ailments associated with impaired	
PT	respiration, has oligo(s) antisense to specific gene(s) or its	
PT	corresponding RNAs, and glucocorticoid or non-glucocorticoid steroid or	
PT	ubiquinone.	
XX		
PS	Disclosure; SEQ ID NO 12061; 872pp; English.	
XX		
CC	The invention relates to a novel pharmaceutical composition, which has a	
CC	first active agent comprising an oligonucleotide antisense to the	
CC	initiation codon, coding region, 5' or 3' end genomic flanking regions,	
CC	5' and 3' intron-exon junctions, or regions within 2-10 nucleotides of	
CC	junctions of genes encoding a polypeptide associated with lung and/or	
CC	nasal airway dysfunction and a second active agent comprising an	
CC	antiinflammatory steroid and ubiquinone. A composition of the invention	
CC	has antiinflammatory, antiallergic, antiasthmatic, hypotensive,	
CC	immunosuppressive, and cytostatic activity. The composition may have a	
CC	use in antisense gene therapy. The composition is useful for treating or	
CC	preventing a respiratory, lung or malignant disease or condition, also	
CC	for enhancing the prophylactic or therapeutic respiratory effect of an	
CC	antiinflammatory steroid in a subject, for reducing or depleting levels	
CC	of, or reducing sensitivity to adenosine, reducing levels of adenosine	
CC	receptor, producing bronchodilation, increasing levels of ubiquinone or	
CC	lung surfactant in a subject's tissue, or treating bronchoconstriction,	
CC	lung inflammation, lung allergies, or a respiratory disease or condition.	
CC	Note: The sequence data for this patent is not represented in the printed	
CC	specification, but was obtained in electronic format directly from WIPO	
CC	at ftp.wipo.int/pub/published_pct_sequences	
XX		
SQ	Sequence 56583 BP; 16610 A; 11169 C; 11927 G; 16877 T; 0 U; 0 Other;	
	Query Match 11.8%; Score 269,097; DB 1; Length 56583;	
	Best Local Similarity 52.7%; Pred. No. 20;	
	Matches 932; Conservative 0; Mismatches 739; Indels 99; Gaps 17;	
QY	298 TCCTGATTTCTACTCTGGCTCATTTTAACTCAGTAGTGAGTGTGGTTCCATAAGT 357	
Db	30070 TCCTGATTTTCATTTTGACCAAGTCATTCGACACAGGTATTTAACTTCCATGAT 30129	
QY	358 TTGTAAGTTTCTCGTTTCTGTTGTTGTTATCTAGATTTAAGCTGTGGTGGT 417	
Db	30130 TTGCAATGGTTTTCGAAGGTTCCCTTTGAAGTTGATT--TCCAGTTTATTCACGTGGT 30186	
QY	418 CAGATAGGACATAGAGTATATTTCAATTGCTTTTATCTGTCGAGACTTGCTTTGTTTT 477	
Db	30187 CTGAGAGAGTGTGTGATAATAATTCCAATTTCTTAAAGTATGGAAGCTCATTTATGGC 30246	
QY	478 GAAATATGATTTCAAATTTGGAG--AGTTTCATAGGTCCTCACAAGAGGTFACAGTCTT 535	
Db	30247 CTATCATATAGTCTATCTTGGAGAAAGTTCCATGATATTGTTGAATAGAAATGTTTATCT 30306	
QY	536 --TGTGTTTTGGTGAATAAGTCGTGAAATATCT-CTAGTCCACTTGTTGTTATGACATCA 592	
Db	30307 GCAGTTGTTGGATGAATATGTTCTGTATATATCTGTTAAGTCCAATTTGTTCCAACTATAG 30366	
QY	593 GTTAGCTCCAGCATTTCTCTGTTTTCGTTTTTGTTCAGATGACCTAACTGTGTGGAGAA 652	
Db	30367 TTTAAATCCATGTTTCTTTTG-TTGACTTTCTGCCTTGAACACTGCTATAGGCTGTGAC 30425	
QY	653 TGGGGTATTGAAGTAGCCCACTATCTGTG-TGAGGTCAAATATGTTATTTAGTGTAG 711	

QY 1777 TAGTGTCTCTGGCTTCTGGATGTTTATGCTGCTGATTAATTTAGACTTAACAATTTCTT 1836
 Db 31507 CATAAATCCAGATTTCTTGGAGGCTTGTTCATATTTCCATATTTCTTTTCTTTGTCTT 31566

QY 1837 TGACCAAGGATTCATTTCTTCTATCTTGTCTTCACTGCTGAGATTTCTTCTATCT 1896
 Db 31567 TGCTGGATGGGTAATTCAT-----AGACCTTAAGCTCTGAATTTCTTTATTCACCT 31620

QY 1897 CTGTATTTCTGCTAGTGAGGCTTGTCTCTGAGGTTCTCTGT-----TGGGTCTTAA 1947
 Db 31621 GTTCAGTTCTATTTCTGAGATTTCCAGAGATTTTCAATTTCTGAAGTGTCCTCAAG 31680

QY 1948 TTTTTCATTTCCAGATTTCTTCTAGTTTGGGTTTGTATTAATTTCTATTTCCACTTT 2007
 Db 31681 TTTCTGAGTTTTCAGTGTGGTTTCTTTAAGCTATCTATTTCCATGAATATTTCTCCCT 31740

QY 2008 CAGTCTCTGAAATGTTTCTACTCATTTTCT 2037
 Db 31741 TCACCTCTGTATCATTTTGTGAATTCCT 31770

RESULT 33
 ID AAA64139/c
 XX AAA64139 standard; DNA; 50000 BP.
 AC AAA64139;
 XX
 DT 20-DEC-2000 (first entry)
 XX
 DE Nucleotide sequence of a beta-tubulin antigen.
 KW Beta-tubulin antigen; inner ear protein; Meniere's disease; autoantibody;
 KW chronic ear disease; autoimmune disease; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200050593-A1.
 XX
 PD 31-AUG-2000.
 XX
 XX 25-FEB-2000; 2000WO-US004795.
 XX
 XX 25-FEB-1999; 99US-0121549P.
 XX
 PA (UYTE-) UNIV TENNESSEE RES CORP.
 XX
 PI Yoo TJ;
 XX
 DR WPI; 2000-558400/51.
 XX
 PT New beta-tubulin antigen in the membranous structure of the inner ear,
 PT reactive with antibodies of patients with Meniere's disease, for
 PT diagnosing Meniere's disease and distinguishing this disease from other
 PT autoimmune ear diseases.
 XX
 PS Claim 3; Page 51-74; 115pp; English.
 CC
 CC The present sequence encodes a beta-tubulin antigen. The protein is an
 CC antigen of the membranous structure of the inner ear protein, and is
 CC reactive with antibodies from patients having Meniere's disease.
 CC Meniere's disease is a chronic ear disease with unknown etiology. Serum
 CC from patients suffering from this disease contain autoantibodies against
 CC a 30 kDa cochlear protein antigen. The disease is believed to be an
 CC autoimmune disease. The beta-tubulin antigen is useful as a target
 CC substance in diagnosing or detecting Meniere's disease and in
 CC distinguishing this disease from other autoimmune ear diseases
 XX
 SQ Sequence 50000 BP; 16814 A; 10211 C; 9526 G; 13449 T; 0 U; 0 Other;
 Query Match 11.8%; Score 268.097; DB 1; Length 50000;
 Best Local Similarity 52.5%; Pred. NO. 21;
 Matches 957; Conservative 0; Mismatches 754; Indels 113; Gaps 18;

QY 298 TCTGATTTCTATCTTGGCTCATTTTAACTCAGTAGTACGTTGTTGGTTCCATAAGT 357
 Db 30648 TCTTGATTTCTATTTGTGAGCCCAATGATCATTCAGGAGCAGGTTATTAATTTCCATGCAT 30589

QY 358 TTGTAAAGTTTCTGTTGTTCTGTTGTTGTTGTTTATCTAGATTTAAGCTGTGGTGGT 417
 Db 30588 TTGCATGGTTTGGAGGTTCTTTTATCATTTGATTTCAAT---TTTATCCACTGTGTT 30532

QY 418 CAGATAGGACATPAGAGTATTAATTCATTTGCTTTTATCTGTCGAGACTTCTTTGTTT 477
 Db 30531 CAGAGAGAGTACTTGCATATACTTCAATTTTCTTAAATTTATTTAGACTTGTTTTGGGC 30472

QY 478 GAAATATGTAATCAATTTTGG--ACAGTTTTCATAGGGTGTGACAAAGAGTACAGTCT- 534
 Db 30471 CTGTCATATGGTCTATCTTTGGTAAATGTTCCAGGTGCTGATGAATAGATGTAATCTG 30412

QY 535 TTGTGTTTGGTGAATAGTCTGTAAATATCTCTAGTCCACTTGGTTTATGACATCAGT 594
 Db 30411 CAGTTGTTGGGTAGAATGTTCTGTAAATATCTGTTAAGTCCATTTCTTTTAGGGTATAGT 30352

QY 595 TAGCTCCAGCATTTCTCTGTTTCGTTTCTTTGTTGAGATGACCTACTGTTGAGAGAATG 654
 Db 30351 TTAAGTCCATTTGTTCTTTGTTGACTTTCTGTCCTTGATGACCTGCTAGTCTGTCAGTG 30292

QY 655 GGTATTTGAAGTAGCCCACTATCTGTGTG-TGAGGTCAATATATGTAATTTTAGCTGTAGCT 713
 Db 30291 GAGTATGAGTCCCTCACTATTTATTTGTTGCTTTCTATCTCATTTCTTAGTCTAGTA 30232

QY 714 GTGCTTGTTTTATGAACCTTGGGTGACATTTGTTGGTGCAATAGACATTAAGAATTCGAA 773
 Db 30231 GTAAATATTTATACATTTGGGAGCTCTAGTGTAGGTACATATATATTTAGGATTTGA 30172

QY 774 TGFCTCTTTGGTGA-TTTTCTTTGATGCCCTATCTAGTATTTCTTCCCAATCTCATCTGC 832
 Db 30171 TATTTCTGTTGGCAAGTCTTTTATCAATTAATATGTCCTCTTTGTTGTTTAAAG 30112

QY 833 TTAGTTTGGGTTTAAAGTCTATTA-GTCAGATATTAATAAGTACTGTAATCGGCTTCTCT 891
 Db 30111 CTGCTATTGCTTTGAAGTTGTTTGTGCTGATATAAGAATACCTACTCTCTGCTCATTTT 30052

QY 892 TAGGGCCATTTGCTTAGAATATCTT-TTCCATCTTTTACTCTAAGGTGATGCTCATCCA 950
 Db 30051 GGTGCTATTTGTCATGGAATATCTTCTCCATCTCCTTTATGTAAGTTAATGTAATCT 29952

QY 951 T---GGTAGGTTGCTTTTGGATGCAGCAGTAGGATGGATTCCTTTGTTTCATATCCATT 1007
 Db 29991 TATGTGTTTGGTAAGTCTCTTGAAGACAGACATATCTTCTGATGAATTTCTATCCATT 29932

QY 1008 CTGTTACCCAGTATCTTTTCTAGAGAATTAAGATCATTTGATCATTTGATGTTGAGAT 1067
 Db 29931 CTGCCATTTCTATCTTTTGAAGTGGAGCATTTAG-----GTTATTTCAATCAATGT 29880

QY 1068 TATCAATGAGCAGTGTGTTGCGATTTCTTTATCTTGCATTTGTGAAGTGTGTGTGTG 1127
 Db 29879 TACANTTGAGATGTGAAGTACTATTCTATTATATGCG-----29842

QY 1128 TG 1187
 Db 29841 -----TAGTTGTGCTGAATACCTTGGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTG 29801

QY 1188 TGTCTTGT 1247
 Db 29800 TGTATATATTTATAGTCTCTGTGAGATTTATGCTTAAAGGGGTTCTATTTTGTGAT 29741

QY 1248 TCATAATTTCTTGAATGTGGTAAACATCTTTAGATTTGAAGTGTGTGTGTGTGTGTGT 1307
 Db 29740 TTGT-----ATGATTTGTTTCAAAATTTAGAGCTACTTTTATCA 29702

QY 1308 AGTCTGCATTTTGAAGATAGATATCTTTACATCTGATTTTATCTTAGAATGCTTTCTT 1367
 Db 29701 GTCTTGTAGTGTGCTGCTTGGTAGTGGCTTAATTCAGCATTTGTCTGAAAAGACTATCTT 29642

```
QY 1368 TCTCCAACTATTGTGACAGAAAGTTTTTCTAAGTCAGTAGTCTGGCCTGACATCTGTAG 1427
|||
Db 29641 TCT--TTCAATTATAAAGCTTAGTTTCACTGGACACAAAATCTTGGCTGATATGTGT 29584
QY 1428 TCTCTTGAGTCTGAGCACATCTGTGAGGCGCTTCTTACATTTTGAGTTTCTTATTGGA 1487
|||
Db 29583 TGTTTAAGGAGGCTAAGATAGAACCCCAATCCCTTCTAGCTTGAGGTTTCTGCT-GA 29525
QY 1488 AAGGTCAGGTGAATCTTAATACATCTGCCCTTATATGTTAAATGGTCTTTTCCCTTG 1547
|||
Db 29524 GAAATCTGCTGTAATCTGATAAGTTTCCCTTTATAGGTTACCTGAT-GCTTTTGTCTCA 29466
QY 1548 CATCTTTTAAATCTCTTTCTTTGTTCTATATCTTTTAGTGATTTGATATTATGCACTGTG 1607
|||
Db 29465 CAGCTCTTAAGATCTTTTCCCTTTGTTGCTGACTTTAGATAAACCCTGATGAAGTGCCTAC 29406
QY 1608 GGGAGTTTCTTTCGGTCCCATCTAATTTGGTGTGTTGATGCTTCTTGACCTTGATAG 1667
|||
Db 29405 ATGATATCTTTTGTGATGAATTTTCCAGGTGTTCTTTCACCTTCTTGTTATTTGGGTAT 29346
QY 1668 GCATCTCTTCTCAAGGTAGGAATTTTCTTTTGGTTTCTTGAATAATTTTCCC 1727
|||
Db 29345 CTAGATCTCTAGCAAGCCAGGAAGTTTCTTGATTTTCTCAACTAGTTTCCCA 29286
QY 1728 TGTCTTTGACCTGCTTCTTCCCTTCTCTATCTTCTTATCTCTTTGGTTT-----TTGCA 1776
|||
Db 29285 AACTTTAGATATCTCTTCTTCTTGGGAACACATAATTAATCTTAGAATCCATCTTAA 29226
QY 1777 TAGTCTCTCGCTTCCTGGATGTTTATGCTGATTTATTTAGATTTATTTAGATTTACTT 1836
|||
Db 29225 CATATCTCAACTTCTCTGGAGCTTGTGTTCAATTTTAAATTTCTGTTTCTTTGACTT 29166
QY 1837 TGACCAAGGTATCCATTTCTTCTATCTTGTCTTCACTGCTGAGATTTCTCTTCTATCT 1896
|||
Db 29165 CTCAGATGGGTAAATTT-GAAGCTTGTCTTCAAGCTCTGAAAGTTCTTCTTTTACTT 29107
QY 1897 CTGTATTTCTCAGTGGGTTGCTCTGAGTTCCTGTGGTCTCTTAATTTTTCAT 1956
|||
Db 29106 ATTCAATTTATATGCTGAGACCTTCCAGTGTATTTTGCATTTCTGTGTCTCTTCATT 29047
QY 1957 TTCCAGATTTCTCTCAGTTTGGGTTTGTGTTTATTAATCTAATTTCCACTTTCAGGCTCTG 2016
|||
Db 29046 TCCAGAGTTT---GATGCTTTTATTTATGCTATTATTTCCCTGT----- 29001
QY 2017 AAATGTTTACTCAATTTTCCCTCCAGTATTTACATTTTCAAGTGTCTTTTAAATGGAAT 2076
|||
Db 29000 AGATTTTCCATTCATATCTGTAACATTTTAAATTTCTTTGAGTGGTATTTCACATTT 28941
QY 2077 ATTCAATTTCTCTCAGGACCTT 2100
|||
Db 28940 CTCTTCTGCTTCTGAGTAGCTT 28917
```

RESULT 34

ABK83575/c

ID ABK83575 standard; cDNA; 136284 BP.

XX ABK83575;

XX AC ABK83575;

XX DT 14-AUG-2002 (first entry)

XX DE Human cDNA differentially expressed in granulocytic cells #146.

XX DE Human; ss; granulocytic cell; DNA chip; bacterial infection;

XX DE fungal infection; parasitic infection; protozoal infection;

XX DE fungal infection; sterile inflammatory disease; psoriasis;

XX DE rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;

XX DE cardiac reperfusion injury; renal reperfusion injury; ARDS;

XX DE adult respiratory distress syndrome; inflammatory bowel disease;

XX DE Crohn's disease; ulcerative colitis; periodontal disease;

XX DE granulocyte activation; chronic inflammation; allergy.

XX OS Homo sapiens.

XX W0200228999-A2.

XX 11-APR-2002.

XX 03-OCT-2001; 2001WO-US030821.

XX 03-OCT-2000; 2000US-0237189P.

XX (GENE-) GENE LOGIC INC.

XX Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;

XX WPI; 2002-435328/46.

XX Detecting granulocyte activation by detecting differential expression of

XX genes associated with granulocyte activation, which serves as diagnostic

XX markers that is useful for monitoring disease states and drug toxicity.

XX Claim 1; SEQ ID NO 146; 114pp; English.

XX The invention relates to detecting (M1) granulocyte (GC) activation

XX (GCA), by detecting the level of expression of gene(s) (Gs) identified by

XX DNA chip analysis as given in the specification, and comparing the

XX expression level to an expression level in an unactivated GC, where

XX differential expression of Gs is indicative of GCA. Also included are

XX modulating (M2) GA by contacting GC with an agent that alters the

XX expression of at least one gene in Gs; (2) screening (M3) for an agent

XX capable of modulating GCA or an inflammation (especially chronic) in a

XX tissue, an allergic response in a subject, exposure of a subject to a

XX pathogen or sterile inflammatory disease using the gene expression

XX profile; (3) detecting (M4) an inflammation (especially chronic) in a

XX tissue, an allergic response in a subject, exposure of a subject to a

XX pathogen or sterile inflammatory disease, by detecting the level of

XX expression of the gene is indicative of inflammation; (4) treating

XX (M5) an inflammation (especially chronic) or in a tissue, an allergic

XX response in a subject, exposure of a subject to a pathogen or sterile

XX inflammatory disease, by contacting a tissue having inflammation with an

XX agent that modulates the expression of gene(s) from Gs in the tissue. M1

XX is useful for detecting GCA; M2 is useful for modulating GCA; M3 is useful

XX for screening an agent capable of modulating GCA preferably in an

XX inflammation in a tissue; M4 is useful for detecting an inflammation

XX (especially chronic) in a tissue, an allergic response in a subject,

XX exposure of a subject to a pathogen or sterile inflammatory disease (e.g.

XX psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,

XX cardiac reperfusion injury, renal reperfusion injury, ARDS, adult

XX respiratory distress syndrome, inflammatory bowel disease, Crohn's

XX disease, ulcerative colitis, periodontal disease; also bacterial

XX infection, viral infection, parasitic infection, protozoal infection,

XX fungal infection and M5 is useful for treating one of the above

XX conditions. The present sequence represents a gene differentially

XX expressed in granulocytes. Note: The sequence data for this patent did

XX not form part of the printed specification, but was obtained in

XX electronic format directly from WIPO at

XX ftp.wipo.int/pub/published_pct_sequences

XX SQ

XX Sequence 136284 BP; 41693 A; 28668 C; 28000 G; 37923 T; 0 U; 0 Other;

XX Query Match 11.7%; Score 267.297; DB 1; Length 136284;

XX Best Local Similarity 51.5%; Pred. No. 12;

XX Matches 915; Conservative 0; Mismatches 777; Indels 83; Gaps 14;

QY 298 TCTTGAATTTCTATCTTGGCTCAATTTTAACTCAGTAGTGTGTGTTGTTCCATAAGT 357

Db 83169 TCTTGAATTTCTGTTTGGACCAATCTGTCGGAGCAGGTTATTTATTTCCATGAT 83110

QY 358 TTGTAAGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 417

Db 83109 TTGATGTTTGAAGGTTCCCTTTGGAGTTGATTTCCAGTTGTTATTCACATGTTGTT-T 83051

QY 418 CAGATAGGACATAGAGTATTTATTTCAATTTGTTCTTTTATCTGTCGAGACTTGTGTTT 477

Db 83050 GAGAGAGGCTTAC---ATAATTTCAATTTTCTTAAATTTATTGAGGTTGTTTATGGC 82994
Qy 478 GAAATATGATTCATTTTGGAG--AGTTTCATAGGCTGTGACAAAGGTACAGTCT- 534
Db 82993 CTATCATATGCTCTATCTTGAGGAGGTTCTATGACCTGTCTAAATGAATGTATTTCTG 82934
Qy 535 TTGTTGTTTGTGAAATAGTCTGTAAATATCTCTAGGTCACCTTGGTTTATGACATCAGT 594
Db 82933 CAGTTGTTGAGTAAATGCTTCTGTATATCTGTAACTCAATTTGTTCCAAAGTATAGT 82874
Qy 595 TAGCTCCAGCATTTCTCTGTTTCTGTTTGTGAGATGACCTTAACCTGTTGGAGAGATG 654
Db 82873 TTAATCCATTTGTTTCTGTTGCTCTCTGTATGACCTGTCTAGTCTGTCTAGTG 82814
Qy 655 GGGTATTGAAGTAGCCCACTATCTGTG-TGAGTCAATATGATTTTATGCTGTAGCT 713
Db 82813 GTGTATTGAAGTCCCACTATTTATTTGGTGTCTGTATCTCAATTTCTTCAGTCTATTA 82754
Qy 714 GTGCTTGTGTTTATGAACCTTGGGTGACATTTGTTTGTGTCATGACATTAAGAAATGCAA 773
Db 82753 GTAACGTTTATAAATTTGGAGCTCCAGTGTAGGTGACATATGTTTAGGACTGTGA 82694
Qy 774 TGTCTCTTGTGGA-TTTTCTTGTGATGCTATGTAGTATTCTCCCAATCTCATCTGC 832
Db 82693 TATTTTCTGTGGACAAAGGCTTTTACATTAATGATTCCTCTTTGTTCTCTTTAA 82634
Qy 833 TAGTTTGGGTTTAAAGTCT-ATTAGTCAGATATTAAATGACTGTATCGGCTTGTCTCT 891
Db 82633 CTGCTGTTGATTAAGTTTGTGTTGCTGTATAGAGTAGTACTCTCTGCTCACITTT 82574
Qy 892 TAGGCCATTTGCTTGAATATC-TTTTCCATCTTTTACTCTAAGGTGATGCTATCCA 950
Db 82573 GGTGTCCATTTGCATGAATGCTTTTCCACCCCTTTACTTTAAGTTTATGTCAGTCT 82514
Qy 951 TGGTAGGT-TGTCITTTTGTGATGAGCAGTAGGATGATCTGTTTTCATATCCATCT 1009
Db 82513 TACATGTTAGGTGAGTCTGAAGGAGCAGATATTGTTGGTGAGTTTATGCAATCT 82454
Qy 1010 GTTACCCAGTATCTTTTCTAGAGAAATTAAGATCAATGAGTCATGATGTTGAGATTA 1069
Db 82453 GCAGTTCTGTATCTTTAAATGGAGCAATTTAGTTATTTACATCAATGTAGTATTA 82394
Qy 1070 TCAATGAGCAGTGTGTTGATCTTGTATCTTGCATCTGAGAGTGTGCTGTG 1129
Db 82393 ATGGAGGTACTGTGCAATGCTATTTGTTGCTGTGCTGCTGTTGTTGTTT 82334
Qy 1130 TGT 1189
Db 82333 GTTTTGTCTTTAACTGTATTTTGTGTTTATAGTCTGTG----- 82292
Qy 1190 TCTGT 1249
Db 82291 -----TCATTTATGCTTTAAAGAGGTATGTATG 82262
Qy 1250 ATATTTCTTGAATGTGGTAACATCTTTAGATTCAGAGTTTCTCTAGGCTCTTTTAG 1309
Db 82261 ATGTGTTTCCAGGATTTGTTCAACATTTAGGTTCTTTTATGTTCTTTGAGTCT 82202
Qy 1310 GTCTGCATTTGAAGATAGATTTCTTTTACATCTGATTTTATCTTAGAATGTCTTTCT 1369
Db 82201 GGTGTTGTTAGTGGCAATTTCTCTGACATGTATGTGTTGTCTGAAAAGACTGTATCT 82142
Qy 1370 TCCAACTATTGTGACAGAAAGTTTCTAAGTGCAAGTGTGCTGGCTCACATCTGTAGTC 1429
Db 82141 TCCCTTCATATGATGCTTAGTTTGTGCTGATATGAATTTCTTGCTGATTAATGTTT 82082
Qy 1430 TCTTGGAGTCTTAGCACATCTGTGAGGCTCTTCTTACATTTTGAGTTTCTATGGAAA 1489
Db 82081 TTTGAGGAGGCTGAAGATAGGCCCAATCCCTTCCAGCTGTGAAGGTTTCTGCT-GAGA 82023
Qy 1490 ACTCAGGTGATTTCTATACATCTGCTTTATATGTTAATGTCCTTTTCCCTTGCA 1549
Db 82022 AATCTGCTGTTAATCTGATAGTGTCTTTCTTTATAGGTACTGTGTGTTGCTCTGCA 81964

Qy 1550 TCTTTTAAATATCTTTCTTTGTTCTATATCTTTTAGTGATTTTATGCACTGTGGG 1609
Db 81963 GCTCTTAGACATATTTCTTTGTTCTTAATCTTTGATAACCTGAAGATAATGCTAGGT 81904
Qy 1610 GAGTTCTTTTCCGGTCCAACTATTTGGTGTGTTGATGCTTTGTTGACCTGTATAGGC 1669
Db 81903 TATGATCTTTTGTGATGAATTTCCAGGTGTCTTTTGTCTCTGTTGTTGATGTCT 81844
Qy 1670 ATCTCTTCTCAAGGTTAGGAAATTTCTTTTGTGTTTCTTGAAAATATTTTCCCTG 1729
Db 81843 AGTCTCTAGCAAGCTGGGAAATTTCTTGAATTTCCCCCAAGTATGTTTCCCAAG 81784
Qy 1730 CTTTGAACCTGCTTCTTCCCT-----TCTCTATTTCTTTGGTTTTTGCATAG 1779
Db 81783 TTTTGAATTTCTTCTTCTCTCAGGAACACCAATTTATTTAGGTTTGGTCAATTAAGCT 81724
Qy 1780 TGTCTCTGCTTCTGATGTTTATGCTGGATTTTATGCTGAATTTTATGACTTAACATTTCTTTGA 1839
Db 81723 AATCTCAGACCTCTTAGATGCTTTGTTGATTTTCTTATCTTTTCTTTGCTGTAT 81664
Qy 1840 CCAAGTATCCATTTCTTCTATCTTGTCTCTCACTGCTGAGATTTCTCTTCTTCTTCT 1899
Db 81663 TGAATTTGGTTAATTTCAAGACCTTTGCTTTGAGCTCTGAATTTCTTCTTCTTCTTCT 81604
Qy 1900 GTATTTCTGTCAGTGGCTTGTCTGAGGTTCTGTTGGTCTTAAATTTTCTTCTTCT 1959
Db 81603 CAGTTCTATTTGCTGAGACTTTCCAGAGCATTTTGCATTTCTGTAAGTATGCTCAATGTT 81544
Qy 1960 CAGATTTCTTCTCAGTTTGGTTTTGTTTATTAAT-----CTATTTTCCATTTT 2008
Db 81543 CCGTAATTTTCTGATGCTTTTCTTTTAAAGCTATTTTCTTCTGAAATTTTCTCCCTT 81484
Qy 2009 AGGTCTCGAAATGTTTACTCATTTTCTTCCCTCCAGT 2043
Db 81483 ACTTCTGTATCTTTTCTTTTATGATTTCTTGTCACT 81449

RESULT 35

AAL05509/c

ID AAL05509 standard; DNA; 32167 BP.

XX AAL05509;

XX AC

XX XX

XX 21-NOV-2001 (first entry)

XX DE

XX Human reproductive system related antigen DNA SEQ ID NO: 8197.

XX Human; reproductive system related antigen; reproductive system disorder;

XX cancer; gene therapy; ds.

XX OS

XX Homo sapiens.

XX PN

XX WO200155320-A2.

XX XX

XX 02-AUG-2001.

XX PF

XX 17-JAN-2001; 2001WO-US001339.

XX PR

XX 31-JAN-2000; 2000US-0179065P.

XX PR

XX 04-FEB-2000; 2000US-0180628P.

XX PR

XX 24-FEB-2000; 2000US-0184664P.

XX PR

XX 02-MAR-2000; 2000US-0186350P.

XX PR

XX 16-MAR-2000; 2000US-0189874P.

XX PR

XX 17-MAR-2000; 2000US-0190076P.

XX PR

XX 18-APR-2000; 2000US-0198123P.

XX PR

XX 19-MAY-2000; 2000US-0205515P.

XX PR

XX 07-JUN-2000; 2000US-0209467P.

XX PR

XX 28-JUN-2000; 2000US-0214886P.

XX PR

XX 30-JUN-2000; 2000US-0215135P.

XX PR

XX 07-JUL-2000; 2000US-0216647P.

XX PR

XX 07-JUL-2000; 2000US-0216800P.

XX PR

XX 11-JUL-2000; 2000US-0217487P.

XX PR

PR	11-JUL-2000;	2000US-0217496P.	PR	01-NOV-2000;	2000US-0244617P.
PR	14-JUL-2000;	2000US-0218290P.	PR	08-NOV-2000;	2000US-0246474P.
PR	26-JUL-2000;	2000US-0220963P.	PR	08-NOV-2000;	2000US-0246475P.
PR	26-JUL-2000;	2000US-0220964P.	PR	08-NOV-2000;	2000US-0246476P.
PR	14-AUG-2000;	2000US-0224518P.	PR	08-NOV-2000;	2000US-0246477P.
PR	14-AUG-2000;	2000US-0224519P.	PR	08-NOV-2000;	2000US-0246478P.
PR	14-AUG-2000;	2000US-0225213P.	PR	08-NOV-2000;	2000US-0246523P.
PR	14-AUG-2000;	2000US-0225214P.	PR	08-NOV-2000;	2000US-0246524P.
PR	14-AUG-2000;	2000US-0225266P.	PR	08-NOV-2000;	2000US-0246525P.
PR	14-AUG-2000;	2000US-0225267P.	PR	08-NOV-2000;	2000US-0246526P.
PR	14-AUG-2000;	2000US-0225268P.	PR	08-NOV-2000;	2000US-0246527P.
PR	14-AUG-2000;	2000US-0225270P.	PR	08-NOV-2000;	2000US-0246528P.
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.
PR	14-AUG-2000;	2000US-0225757P.	PR	08-NOV-2000;	2000US-0246609P.
PR	14-AUG-2000;	2000US-0225758P.	PR	08-NOV-2000;	2000US-0246610P.
PR	14-AUG-2000;	2000US-0225759P.	PR	08-NOV-2000;	2000US-0246611P.
PR	18-AUG-2000;	2000US-0226279P.	PR	08-NOV-2000;	2000US-0246613P.
PR	22-AUG-2000;	2000US-0226681P.	PR	17-NOV-2000;	2000US-0249207P.
PR	22-AUG-2000;	2000US-0226682P.	PR	17-NOV-2000;	2000US-0249208P.
PR	22-AUG-2000;	2000US-0227182P.	PR	17-NOV-2000;	2000US-0249209P.
PR	23-AUG-2000;	2000US-0227182P.	PR	17-NOV-2000;	2000US-0249210P.
PR	30-AUG-2000;	2000US-0228924P.	PR	17-NOV-2000;	2000US-0249211P.
PR	01-SEP-2000;	2000US-0229287P.	PR	17-NOV-2000;	2000US-0249212P.
PR	01-SEP-2000;	2000US-0229343P.	PR	17-NOV-2000;	2000US-0249213P.
PR	01-SEP-2000;	2000US-0229344P.	PR	17-NOV-2000;	2000US-0249214P.
PR	01-SEP-2000;	2000US-0229345P.	PR	17-NOV-2000;	2000US-0249215P.
PR	05-SEP-2000;	2000US-0229509P.	PR	17-NOV-2000;	2000US-0249216P.
PR	05-SEP-2000;	2000US-0229513P.	PR	17-NOV-2000;	2000US-0249217P.
PR	06-SEP-2000;	2000US-0230437P.	PR	17-NOV-2000;	2000US-0249218P.
PR	06-SEP-2000;	2000US-0230438P.	PR	17-NOV-2000;	2000US-0249244P.
PR	08-SEP-2000;	2000US-0231242P.	PR	17-NOV-2000;	2000US-0249245P.
PR	08-SEP-2000;	2000US-0231243P.	PR	17-NOV-2000;	2000US-0249264P.
PR	08-SEP-2000;	2000US-0231244P.	PR	17-NOV-2000;	2000US-0249265P.
PR	08-SEP-2000;	2000US-0231413P.	PR	17-NOV-2000;	2000US-0249297P.
PR	08-SEP-2000;	2000US-0231414P.	PR	17-NOV-2000;	2000US-0249298P.
PR	08-SEP-2000;	2000US-0232080P.	PR	17-NOV-2000;	2000US-0249300P.
PR	08-SEP-2000;	2000US-0232081P.	PR	01-DEC-2000;	2000US-0250160P.
PR	12-SEP-2000;	2000US-0231958P.	PR	01-DEC-2000;	2000US-0250391P.
PR	14-SEP-2000;	2000US-0232397P.	PR	05-DEC-2000;	2000US-0251030P.
PR	14-SEP-2000;	2000US-0232398P.	PR	05-DEC-2000;	2000US-0251988P.
PR	14-SEP-2000;	2000US-0232399P.	PR	05-DEC-2000;	2000US-0256719P.
PR	14-SEP-2000;	2000US-0232400P.	PR	06-DEC-2000;	2000US-0251479P.
PR	14-SEP-2000;	2000US-0232401P.	PR	08-DEC-2000;	2000US-0251856P.
PR	14-SEP-2000;	2000US-0233063P.	PR	08-DEC-2000;	2000US-0251868P.
PR	14-SEP-2000;	2000US-0233064P.	PR	08-DEC-2000;	2000US-0251869P.
PR	14-SEP-2000;	2000US-0233065P.	PR	08-DEC-2000;	2000US-0251989P.
PR	21-SEP-2000;	2000US-0234223P.	PR	08-DEC-2000;	2000US-0251990P.
PR	21-SEP-2000;	2000US-0234274P.	PR	11-DEC-2000;	2000US-0254037P.
PR	25-SEP-2000;	2000US-0234997P.	PR	05-JAN-2001;	2001US-0

PN WO200295031-A1.
 XX 28-NOV-2002.
 PD 10-MAY-2002; 2002WO-ES000221.
 XX 18-MAY-2001; 2001ES-00001133.
 PR (CNSJ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.
 XX Montoliu Jose L, Giraldo Carballo P, Busturia Jimeno AM;
 XX WPI; 2003-112151/10.
 DR New DNA segment with insulating activity, useful for reducing the
 PT variation of transgene expression, in animal cells, caused by integration
 PT at different genomic locations.
 XX Claim 1; Page 54-55; 57pp; Spanish.
 XX The invention relates to a 3711 bp region of the mouse tyrosinase gene
 CC referred to as the locus control region (LCR; ABV76777), and its
 CC functional fragments and analogues (ABV76778-ABV76781). The tyrosinase
 CC gene from which the LCR is derived is located on chromosome 7 and plays a
 CC role in pigmentation. A portion of the full-length LCR sequence,
 CC designated HS (ABV76779) was identified as a DNaseI hypersensitive region
 CC that was found to be deleted in the mutant tyrosinase allele chinchilla-
 CC mottled, which confers a variegated pigmentation phenotype. The full-
 CC length LCR, or its functional fragments or analogues may be incorporated
 CC into animal cell expression constructs to protect against the effects of
 CC transgene integration at different locations in the genome. The LCR-
 CC derived sequence reduces the variability of transgene expression and
 CC associated with integration at different chromosomal locations, and
 CC ensures optimal expression (as regards time, place and quantity),
 CC regardless of the site of integration. The present sequence represents
 CC the full-length LCR of the mouse tyrosinase gene. In addition to its
 CC protective effect against variable transgene expression, this sequence
 CC also contains an enhancer element (denoted as Box AB in the
 CC specification); however, this enhancer is not required for protection
 CC against transgene expression variability
 XX SQ Sequence 3711 BP; 1037 A; 595 C; 650 G; 1429 T; 0 U; 0 Other;

Query Match 11.0%; Score 249.899; DB 1; Length 3711;
 Best Local Similarity 62.2%; Pred No. 89;
 Matches 574; Conservative 0; Mismatches 296; Indels 53; Gaps 12;

QY 292 TCTATTTCTGATTTCTATCTGGCTCAATTTTAACTCAGTAGTCAGTGTGTTGGTTCC 351
 DB 617 TTCTTTCTTTATTTCTGCTTACCTCAAGTATCCAGTAGGGAGTGTGCAGCTTC 676
 QY 352 ATAAGTTTGAAGTTTCTGTTGTTCTGTTCTGTTCTTATCTAGATTAAAGCTGT 411
 DB 677 ATGAGTTTGTAGTTTCTGTTGTTGA-----ATCCACCTTTACTCCA 720
 QY 412 GGTGTGTAGATAGGACATAGAGTATTATTCAATTTGTTCTTTTATCTGCGAGCTTGGT 471
 DB 721 TGTGTCTTAATAAGACACAGGCAATTTTCACTTTCTTATCTGTTGAGGCTTATT 780
 QY 472 TGTTTTGAATATGATTCAATTTTGGAG-AGTTTCATAGGGTCTGACAGAGGTACA 530
 DB 781 TGTGAACAACTATATGATCAATTTTGGAGAAAGTTAAGGATTGCTGAGAAAGGTGCC 840
 QY 531 GTCTTTTGTGTTTGGTGAATAGTCTGTAATATCTCTAGGTCCACTTGGTTTATGACAT 590
 DB 841 TCTTTTGTGTTTGGTGAATAGTCTGTAATATCTCTAGGTCCACTTGGTTTATGACAT 590
 QY 591 CAGTTAGTCCAGCAATTTCTGTTGTTTCTGTTTCTGTTGAGATGACCTAACCTGTTGGAG 650
 DB 887 CTTTACCTCCCAATTTCTTAGTTTGA-TTTTTGTCTAGATGACATGCTATTTCATGAG 945
 QY 651 AATGGGGTATTCAAGTAGCCCACTATCTGTGTGTGAG-GTCAATATGTGATTTTAGCTGT 709

DB 946 AGTAGATATTGAAGTCTCCCTACTATTAAATGATTAGATTCAATTTGTGATTGAGCAAT 1005
 QY 710 AGCTGTGCTTTTATGAACCTTGGTGACATTTGTTTGGTGCATAGACATTAAGAATT 769
 DB 1006 AGTAATGTTTCACTT-CCAAATGTGTACCTTGTCTTTGGGACATAGATGTTAAGACT 1064
 QY 770 GCAATGTCCTCTTGG-TGGATTTTCTTTGATGCCCTATGATATCTTCCCAATCTCAT 828
 DB 1065 GAAGTGTATCATGGCTAATTTTCTTTGATGAGTATGAAATGTCATTTTCCATCTCTA 1124
 QY 829 CTGCTTAGTTTGGGTTTAAAGTCT-ATTAGTCAGATATTAAATGACTGTATCGGCTTGC 887
 DB 1125 TTGACTTAATTTCTGTAAGTGTGATTTGCTAGATATTAGATAGCTAAACAGCTTGC 1184
 QY 888 TTCTTAGGCCCAATTTGCTTTAGAAATCTTTTCCATCTTTTACTCTAAGGTGATGCTAT 947
 DB 1185 TTCTTGAGTGCACTGCTTGGAAATATTTTTCACACCTTAACCTGGAGGTATTGCTAT 1244
 QY 948 C--CATGGTAGTGTCTTTTGGATGCAGCAGTAGGATGATCTTGTTCATATCCA 1005
 DB 1245 CTGATGTTGAGTCTATTTTGTACTCAGTGAATGATGATCTCTTTTGCATCCA 1304
 QY 1006 TTCTGTATCCCAATCTTTTCTAGAGAAATTAAGATCAATTCAGTCAATGATGTTGAGA 1065
 DB 1305 CTCTGTTTCCCTGTGTGTTT-----ATTGAGTAATTCAGTCCATTTACTTTGAGA 1355
 QY 1066 ATTATCAATGACAGTGTGTTGGATCTTGTATCTTTCACCTTGTGAGTGTGTGTG 1125
 DB 1356 AATATTAAATGATCAATGATCTTCATTTCTGTTATTTG-----GTGGTGGTGTGATAG 1410
 QY 1126 TGT 1184
 DB 1411 TAGTGATAGTTGTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1470
 QY 1185 CTGTGTCTTGT 1207
 DB 1471 CTGTATCATGATTCCTATATGC 1493

RESULT 40

AAC64370

ID AAC64370 standard; DNA; 125910 BP.

XX AC AAC64370;

XX DT 07-FEB-2001 (first entry)

XX DE Human KCNQ5 (KCN6q) gene sequence SEQ ID NO:1.

XX KW Human; KCNQ5; KCN6q; chromosome 6; voltage-gated potassium channel;

XX KW Stargardt-like macular dystrophy; cone-rod macular dystrophy;

XX KW Salla disease; ophthalmological; auditory; central nervous system;

XX KW cardioactive; anticonvulsant; gastrointestinal; muscular active;

XX KW age-related macular degeneration; macular degeneration; deafness;

XX KW epilepsy; neuropsychiatric disorder; heart disorder; muscle disorder;

XX KW gastrointestinal disorder; ds.

XX OS Homo sapiens.

XX XX WO200061606-A1.

XX XX 19-OCT-2000.

XX XX 10-APR-2000; 2000WO-US009587.

XX XX 14-APR-1999; 99US-0129274P.

XX XX (MERI) MERCK & CO INC.

XX XX Petrukhin K, Caskey CT, Li W, Metzker ML;

XX XX WPI; 2000-647417/62.

XX XX P-PSDB; AAB24241.

XX Voltage-gated potassium channel KCNQ5 DNA and protein, for identifying
PT inhibitors and activators which can treat e.g. Stargardt-like macular
PT dystrophy, cone-rod dystrophy, Salla disease, deafness, and epilepsy.
XX Claim 3; Fig 1; 99pp; English.
XX The present sequence represents the human KCNQ5 (also called KCN6q) gene,
CC which encodes a voltage-gated potassium channel protein. Human KCNQ5 has
CC ophthalmological, auditory, central nervous system (CNS), cardioactive,
CC anticonvulsant, gastrointestinal and muscular active activities.
CC Sequences and methods from the present invention are useful for
CC identifying activators or inhibitors of KCNQ5 protein. These activators
CC and inhibitors are useful for treating Stargardt-like macular dystrophy,
CC cone-rod dystrophy, Salla disease, age-related macular degeneration,
CC other forms of macular degeneration, deafness, epilepsy, and different
CC forms of neuropsychiatric, heart, gastrointestinal, and muscle disorders.
CC Stargardt-like macular dystrophy and cone-rod dystrophies are located at
CC chromosome 6q
XX
SQ Sequence 125910 BP; 40132 A; 24180 C; 23166 G; 38360 T; 0 U; 72 Other;
Query Match 11.0%; Score 249.898; DB 1; Length 125910;
Best Local Similarity 51.0%; Pred. No. 15;
Matches 878; Conservative 0; Mismatches 766; Indels 77; Gaps 13;

Qy 298 TCCTGATTTCTATCTTGGCTCAATTTTAACTCAGTAGTAGTGGTTGGTTCCATAGT 357
Db 65416 TCCTGATTTCTATCTTGGCTCAATTTTAACTCAGTAGTAGTGGTTGGTTCCATAGT 65475

Qy 358 TTGTAAGTTTCTCTGTTCTGTTGTTGTTGTTATCTAGATTTAAAGCTGGTGGT 417
Db 65476 TTGCAATGGCTTGAAGTTCATTTGGAGTTGATTTTC---AGTTTATTCACCTGGT 65532

Qy 418 CAGATAGACATAGATATATTTCAATGTCCTTTTATCTGTCGAGACTGCTTGGTTT 477
Db 65533 CTGAGAGAGTCTTGGATATATTTCAATTTCTTAAATTTATTCAGGCTCATTTTGGC 65592

Qy 478 GAAATATGATTCATTTTGGAGTTTCATAGGGTCTGACAGAGGTACAGTCTTTC 537
Db 65593 CTGTCGATGGTTCACCTTGGAGAAATTCATGCACTGAATGAATGATATTTCTATGG 65652

Qy 538 TGTTTGGTGAATAGTCTGTAAATATCTCTAGTCCACTTGGTTTATGACATCAGTTAG 597
Db 65653 TTGTTGATGGAATGTTCTGTATATATCTGTAACTATTTGTTTCCAGGTATAGTTA 65712

Qy 598 CTCAGCATTTCTCTGTTTCTGTTTGTGAGATGACCTAACTGTTGGAGAGAAATGGGG 657
Db 65713 AATCCATGTTTCAATGTTGACCTTCACTGCTTGAATATCTGTCAGTGTGTCAGTGGAG 65772

Qy 658 TAATGAAGTAGCCCACTATCTGTG-TGAGGTCAATATGTGATTTTATGCTGTAGCTGTG 716
Db 65773 TAATGAAGTCTCCCACTATATTAATGTTGCCACTGCTCATTTCTAAGTCTGTAGTA 65832

Qy 717 CTCTTTTATGAATCTGGGTGACATGTTGTTGGTGATAGACATTAAGAAATGCAATGT 776
Db 65833 ATTGTTTTATAAATTTGATAGTTCAGTGTTCAGTGCATATAGCTTTAGGATTTGTATGT 65892

Qy 777 CCTCTGGTGATTTTCTTTGAGCCTATGATATTTCTCCCAATCTCATCTGCTTAG 836
Db 65893 TTTCCTGTGACATGACCTTTATCAATATATATAATGCCCTCTTTGTCITTTTAAATGC 65952

Qy 837 TTTTGGTTTAAAGTCTATAGTCAGATATTAATAATGACTGTATCGGCTTCTCTTAGGG 896
Db 65953 TGTTCCTTAAAGTTTGTGTC-----TAAGATAGGTACTCCCGTTTCTTTTGGTGT 66007

Qy 897 CCATTTCCTTAGATATC-TTTTCCATCCCTTTACTCTAAGTGAATGCTATCAT---G 952
Db 66008 CCATATTCCTAAAGATGCTTTTCCACACCTTTACCTTAAAGTTGTGAGACTCCCTATGT 66067

Qy 953 GTAGGTGCTTTTGGATGACAGTAGATGGATCTTTGTTTTCATATCCATCTCGTT 1012
Db 66068 GTTAGCGAGTCTCTTGAAGTCAGCAGGTAGTTGGTTGGTGAATCTTATCAATCTGCA 66127

RESULT 41
AAS63340
ID AAS63340

standard; DNA; 6207 BP.

Qy 1013 ACCAGTATCTTTTCTAGAGAAATTAAGATCAATGAGTCATTGATGTTGAGATTTATCA 1072
Db 66128 ATTCTATATCTTTT-----AAGTGGAGCATTTAAGCCATTTCAATTTCAAGTTAGTA 66179

Qy 1073 ATGAGCAGTGTGTTGGGATCTTGTATCTTGCACCTGTGGAAGTGTGTGTGTGTGT 1132
Db 66180 TTGAGATGTGAGGTACCAATTCATTCATATGC-----TATTTGTTCCCTGTATACCT 66232

Qy 1133 GT 1192
Db 66233 TGGTTTTTGTGTTTTTGTGTTTTTAAATTTGTTTATTTTATAGTCTGTGTA 66292

Qy 1193 TGT 1252
Db 66293 GATTTATGCTTTAAAGAGAGTGTCTGTTTGAAGTGTTCAGGATTTGTTTCAAGAT--- 66348

Qy 1253 TTTTCTTGAATGTGGGTAACATCTTTAGATTCAGATTTTTCCTAGCCTTCTTAGGTC 1312
Db 66349 -----TTAGAGCTCTCTTTTAGCAGTTCTTGTAGTGG 66379

Qy 1313 TGCATTTGAAGATAGATATTTCTTTACATCTGATTTTATCTTAGAATGCTTTCTTTCTCC 1372
Db 66380 TGGCTTGGTAGTGGTGAATTTCTCTCAGAAATTTGTCTGCTGAAAGAGCTGATCTTTCC 66439

Qy 1373 AACTATTGTGACAGAAAGTTTTTCTAAGTGCAGTAGTCTGGCCTGACATCTGTAGTCTCT 1432
Db 66440 TTCTATATGAAGCATAGTTTGTGTTGGATACAAAATTTCTGGCTGGTAAATCTTTTGCTT 66499

Qy 1433 TGGAGTCTGTACACATCTGTGCAGGCTTCTTACATTTTGTGATTTTCTATTGGAAGT 1492
Db 66500 GAGGAGGTGAAGACAGCCCAATACCTTCTAGCTATAGGGTTCTGCT-GAGAAAT 66558

Qy 1493 CAGGTGATATTTCTAATACATCTGCCTTTATATGTTAAATGGTCTTTTCCCTGCACT 1552
Db 66559 CTGCTGTTAACTAATAGAGTGTGTTTATTTATAGGTACCAAGTACTTTGT-CTCACTGCT 66617

Qy 1553 TTTAATATCTTTCTTGTGTTCTATCTTTAGTGTATGATTTATGCACTGTGCGGAG 1612
Db 66618 CTAAATTTCTTTCTTGTGTTCTAATTTAGATTTACCTGATGACAAATGCTAGTGTAT 66677

Qy 1613 TTTCTTTCCGGTCCAATCTATTTGGTGTGTTTGTATGCTTTGTACCTTGTATAGGCATC 1672
Db 66678 GATCTTTTGTCAATGAATTTTCCAGGTGTTCTTTGCTCTTGTATTGGAATGCTAGG 66737

Qy 1673 TCTTTCTCAGGTAGGAAATTTTCTTTTGTGTTTCTTGTGAAATATTTTCCCTGCTT 1732
Db 66738 TCTTAGCAATACTGGGAAAGTTTTCCTCTGTATTTCCTGAAATATGTTTCCAAACTT 66797

Qy 1733 TTGACCTGCTTCTTCCCT-----TCCCTCTATTCTCTTGGTTTGTGCTAGTG 1781
Db 66798 TTAGATTTCTCTTCTTCTCAAGACACTGATTTATTTAGCTTTGGTCACTAACATAA 66857

Qy 1782 TCTCTGGCTTCTGATGTTTATGCTGGATTAATTTAGACTTAACTTTTCTTTTGACC 1841
Db 66858 TCCAGACTTCTTGGAGCTGTGTTTCTTATTTCTTATTTCTTTTCTTTTCTTTTCTTTT 66916

Qy 1842 AAGGTATCCATTTCTTCTATCTTGTCTTCTGCTGCTGAGATTTCTCTTTCTATCTTGT 1901
Db 66917 GAGTGGTTAACTGAAGACCTTGTCTTAGACTCTGAATCTTTCTTTCTTTCTTTCTTT 66976

Qy 1902 ATTCTGTAGTGGCTTGTCTCTGAGTTCTGTGGTGTCTTAAATTTTCTTTTCTTTTCCA 1961
Db 66977 ATTCTATTGCTGAGACT--TTCAGAGCATTTTGCATTTCTATAGGGTGTCCAAAGTTT 67034

Qy 1962 GATTTCTTCAAGTTTGGGTTTGTATTATTAATCTATTTCC 2002
Db 67035 ACTGAAGTTTGACTGTTTATTATTATGCTATCTATTTC 67075

QY 586 GACATCAGTTAGCTCCAGCATTTCTCTGTTTCGTTTTTTTGTGTGAGTACCTTAAGTCTTG 645
DB 21320 GGTATAGTTTAAAGTCCATTTTCTCTTG-TGACATTTCTGCTTGAAGACCTGCTAGTG 21378
QY 646 GAGAGAAATGGGGTATTGAAGTAGCCACTATCTGTGTG-TGAGGTCAATATGTGATTTTA 704
DB 21379 CTGCCAGTGGAGTATGAAGTTCGCACTATTTTATGTTGCCATCTAGCTCAGTCTCTTA 21438
QY 705 GCTGAGCTGCTGCTGTTTTATGAACCTGGGTGACATGTGTTGGTGGCATAGACATAA 764
DB 21439 GGTGAGTAGTGAAGGTTTTATAAATTTGTGAGCTCCAGTGTAGGTTAATATATATTA 21498
QY 765 GAATGGCAATGCTCTCTGGTGG- - - - -TTTTCTTGATGCTATGTAGTATTC 815
DB 21499 GGATTTGACATTTTCTGTTGGACTAATCTTTTAAATATATATAGCCACTTTGT 21558
QY 816 TTCCCAATCTCATCTGCTTGTGTTGGGTTTAAAGTCTATATAGTCAGATATTAATAAGCT 875
DB 21559 CTTTCTTTTTTAACTGCTGTGCTCTGAAGTCTGTTTGAAGTGGTATAGACTAGC 21618
QY 876 GTATCGGCTTCTCTTAGGCGCAATTTGCTTAGAATATCTTTCCATCCCTTTTACTCTAA 935
DB 21619 TACTTCTGCTTGTGTTTCCATTTGCGTGGAAATATTTTCCACCCCTTTTACT-TAA 21677
QY 936 GGTGATGCTATCCATG- - -GTAGTTGCTTTTTTGGATGACAGCAGTAGGATGGATCTT 992
DB 21678 GTTTATATGAATCCTCATGTGTCAGGTGAGTCTCTTGAAGCAGAGATACTTGGTTGGT 21737
QY 993 GTTTTCATATCATCTGTTTACCAGTATCTTTTCTAGAGAAATTAAGATCATTTGAGTC 1052
DB 21738 GGATTTTATCCATCTGCCATCTGCACTCTCTTAAAGTGGAGCAATTTAGA- - - - -TC 21789
QY 1053 ATTGATGTTGAGAAATTATCAATAGCAGTGTGTTGGATTCCTGTTATCTTGCACTGTG 1112
DB 21790 ATTTACGTTTAAATGTTAGTATTTAGATTTGAGTACTGTTCAATTAATCATGCTA- - - - - 21844
QY 1113 AAGT 1172
DB 21845 - - - - -GTGTGCTC 21852
QY 1173 GT 1232
DB 21853 CTTAAATCTTTTCCCATTTGTTATTTTATAGGCTTGTGATTTTATGCTTTA 21912
QY 1233 GGAATTTATTTATTTATTTCTTGAATGTGGTAACATCTTTAGATTTGAAGTTT 1292
DB 21913 AGGAGTTCTATTTTGGTGTATTT- - - - -TGAGTTTCTTCCAAATTTGGAACCT 21966
QY 1293 CTCCTAGCTCTTTAGTCTGCATTTGAAGATAGATTTCTTTACATCTGATTTATCT 1352
DB 21967 CTTACTACTCTGTGAGT 22026
QY 1353 TAGAATG- - -TCTTTCTTCTCCAACTATTGTGACAGAAAGTTTTTCT- - -AAGTGCAGTA 1407
DB 22027 GAAAGGAATTTTAACTCTCTCTCATTTATGAAGCTTAGTTTGTCTGATATAAAAAA 22086
QY 1408 GTCTGGCTGACATCTGTAGTCTCTTGGAGTCTGTAGCACAATCTGTGAGGCGCTCTTA 1467
DB 22087 TTTGTGGCTGACAATTAATTTGTTTTCAGGAGGCTAAAGATGGGACCCCAATACATCTGC 22146
QY 1468 CATTTTGTAGTTTCTATTGGAAAAGTCAGGTGTAATTTCTAATACATCTGCCCTTATATGTT 1527
DB 22147 CTTGTAAGGTTTCTGCTTGA- - - - -AATAGGTTTTCTTTATAGATT 22189
QY 1528 AATTGGCTTTTTTCCCTTGACATCTTTTAAATCTTTTGTGCTATACCTTTTAGTGA 1587
DB 22190 AACTGATGCTTTGTTCCATAGCTCTTAAAGATTTTTTCTGCTGTCTTTGACTTTAGACAA 22249
QY 1588 TTTGATTTATGCACTGTGGGAGTTCTTTTCCGGTCCCAATCTATTTGTGTGTTTT-GT 1646
DB 22250 CTTGATCTGATGCTGTACATGATAATCTTTTTTGTGATGAATTTCCAGTGTGTTTTCT 22309
QY 1647 ATGCTTCTTGTACCTTGTATGGCATCTCTTTCTCAAGGTTAGGAAATTTTTTCTTTTTTGG 1706

DB 22310 AAGCTTCTGTATTTGAATGTCTAGATCTCTAGCAAGCTGGGAGCTTTCCAGTT- - - 22366
QY 1707 TTTTCTGAAAAATATTTCCCTGCTTTTGACCTGCTTCTTCCCTTCTCTATTTCCCTTT 1766
DB 22367 GTTTCCTCAATAAGTTTTTCCAACTTTTAGATTTCTCTTTATTTCTGGGAATACCAATT 22426
QY 1767 GGTTTTGCATAGTCTCTGCTGCTTCCCTGGAGTGTTTATGCTGCTGATTTT- - - - - 1819
DB 22427 ATTTTGTGTTGGCCATTTAAACATTAATTTCAATTTCTGTGCTTTTGTTCATTTAC 22486
QY 1820 - - - - -AGACTTAACATTTCTTTGACCAAGGTATCCATTTCTTCTATCTGTGTTTCCACTGC 1875
DB 22487 TCAATTTTTTAAATTTGCTTGTCTGATTTGGTTAAATTTCAAAACCTTGTCTCGAGCC 22546
QY 1876 CTGAGATTTCTCTCTCTATCTCTGTTATCTGCTAGTGAGGCTTCTCTGAGGTTCTCTG 1935
DB 22547 TTGAAGTTTTTCTCTCTACTGTTTGTAGTCTATGCTGACACTTCTCTAGTGCATTTGTG 22606
QY 1936 TTGGGTTCTTAATTTTTTCAATTTCCAGA- - - - -TTTCCCTTCAGTTTGGTTTGTATTA 1991
DB 22607 TTTCTTAGCGTATTTTTCATTTCCAAAATATGATTTGTTCTGCGTTGATCTG 22666
QY 1992 ATTTCTATTTCCACTTTCAGGCTCTGAAATGTTTACTCATTTTCTCCCTCCAGTATTTACAT 2051
DB 22667 TTTCTCTGATAATTTTTCATCCATATCTCTGATTTGTTTTTAAATTCATTCAGTTGTT 22726
QY 2052 TTTTCATAGTCTTCTTAAT 2070
DB 22727 TTTACCTTTCTCTGAT 22745
RESULT 43
AAC64370/c
ID AAC64370 standard; DNA; 125910 BP.
XX AAC64370;
XX 07-FEB-2001 (first entry)
XX Human KCNQ5 (KCN6q) gene sequence SEQ ID NO:1.
DE Human; KCNQ5; chromosome 6; voltage-gated potassium channel;
KW Stargardt-like macular dystrophy; cone-rod macular dystrophy;
KW Salla disease; ophthalmological; auditory; central nervous system;
KW cardiolipin; anticonvulsant; gastrointestinal; muscular active;
KW age-related macular degeneration; macular degeneration; deafness;
KW epilepsy; neuropsychiatric disorder; heart disorder; muscle disorder;
KW gastrointestinal disorder; ds.
OS Homo sapiens.
XX WC200061606-A1.
XX 19-OCT-2000.
XX 10-APR-2000; 2000WO-US009587.
XX 14-APR-1999; 99US-0129274P.
XX (MERI) MERCK & CO INC.
XX Petrukhin K, Caskey CT, Li W, Metzker ML;
XX WPI; 2000-647417/62.
XX P-PSDB; AAB24241.
XX Voltage-gated potassium channel KCNQ5 DNA and protein, for identifying
XX inhibitors and activators which can treat e.g. Stargardt-like macular
XX dystrophy, cone-rod dystrophy, Salla disease, deafness, and epilepsy.
PS Claim 3; Fig 1; 99pp; English.
XX

Db 17726 G 17726

RESULT 45
ADA02777
ID ADA02777 standard; DNA; 96595 BP.
XX
AC ADA02777;
XX
DT 06-NOV-2003 (first entry)
XX
DE Mouse Ptpk carcinoma associated gene, SEQ ID NO:1295.
XX
KW Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer; breast;
KW prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
KW gene; ds.
XX
OS Mus sp.
XX
PN WO2003057146-A2.
XX
PD 17-JUL-2003.
XX
PF 26-DEC-2002; 2002WO-US041414.
XX
PR 26-DEC-2001; 2001US-00035832.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW;
XX
DR WPI; 2003-587068/55.
XX
PT New recombinant nucleic acid encoding carcinoma associated protein,
PT useful for preparing compositions for treating carcinomas.
XX
PS Claim 1; SEQ ID NO 1295; 245pp; English.
XX
CC The invention relates to recombinant carcinoma associated (CA) nucleic
CC acid sequences from mouse and human (ADA01482-ADA03094), and to
CC recombinant carcinoma associated proteins (CAP) encoded by them. The
CC invention also encompasses expression vectors and host cells comprising a
CC CA nucleic acid, a polypeptide (especially an antibody) that specifically
CC binds to the protein, and a biochip comprising CA nucleic acid or
CC fragments thereof. The sequences of the invention were identified using
CC oncogenic retroviruses, which insert into the genome of the host organism
CC at random. Many of these do not carry transduced host oncogenes or
CC pathogenic trans-acting viral genes, meaning that cancer incidence is a
CC direct consequence of the effects of proviral integration into host
CC protooncogenes. The CA nucleic acid sequences can be used to diagnose
CC carcinoma (especially breast cancer, prostate cancer, lymphoma or
CC leukaemia) or a propensity to carcinoma by determination of the sequence
CC of a CA gene, or by determination of CA gene expression in particular
CC tissues. CA nucleic acids, proteins and antibodies are also useful as
CC therapeutic agents and in screening and evaluating drug candidates. The
CC present sequence represents a specifically claimed murine CA nucleic acid
CC sequence of the invention. Note: The complete sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 96595 BP; 27994 A; 18068 C; 18732 G; 30838 T; 0 U; 963 Other;

Query Match 3.1%; Score 71.4; DB 1; Length 96595;
Best Local Similarity 74.4%; Pred. No. 36;
Matches 90; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 1069 ATCAATGAGCAGTGTTCGTGATTCCTGTTATCTTCGACTTCGAAAGTGTGTGTGTGT 1128
DB 17606 AACTATGACTCATTTCTGTGATATCTAGATGCTTTTGTGTGTGTGTGT 17665
QY 1129 GT 1188

Db 17666 GTATGT 17725
QY 1189 G 1189
Db 17726 G 17726

Search completed: August 25, 2004, 07:09:04
Job time : 2296 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2004, 07:11:15 ; Search time 3651 Seconds
(without alignments)
3,978 Million cell updates/sec

Title: US-10-664-775-4

Perfect score: 2279
Sequence: 1 gatcaactcctctagtgaaag.....ttgtaattctaggtgctgat 2279

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 0.5

Searched: 39 seqs, 3186700 residues

Total number of hits satisfying chosen parameters: 78

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : rni4.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	588.3	25.8	29604	1	US-08-781-891-207
2	588.3	25.8	29604	1	Sequence 207, App
C	3 565.399	24.8	51259	1	US-08-781-891-209
C	4 565.399	24.8	51259	1	Sequence 209, App
5	521.797	22.9	51259	1	US-08-781-891-209
6	521.797	22.9	51259	1	Sequence 209, App
7	323.899	14.2	16442	1	US-09-618-166-208
8	323.899	14.2	16442	1	Sequence 208, App
9	253.097	11.1	193303	1	US-09-497-855A-37
10	253.097	11.1	193303	1	Sequence 37, Appl
C	11 252.997	11.1	392000	1	US-09-497-855A-44
12	251.899	11.1	10409	1	US-10-027-983-11
13	248.997	10.9	168575	1	US-08-772-440-33
14	242.997	10.7	116592	1	US-09-426-230-1
15	239.897	10.5	786431	1	US-09-818-512-3
16	239.797	10.5	32042	1	US-09-751-389-3
17	239.797	10.5	32042	1	Sequence 44, Appl
18	227.997	10.0	40328	1	US-09-245-281-44
19	208.398	9.1	202001	1	US-09-340-620A-63
C	20 204.799	9.0	99916	1	US-08-742-185-102
C	21 200.5	8.8	148567	1	US-09-734-674-3
C	22 200.5	8.8	148567	1	US-09-816-095-3
C	23 191.799	8.4	168575	1	US-09-801-876B-3
C	24 186.899	8.2	2116	1	US-10-254-869-3
C	25 186.899	8.2	2116	1	US-09-426-230-1
26	186.1	8.2	392000	1	US-09-377-437-11
27	177.9	7.8	99916	1	US-09-377-437-67
C	28 176	7.7	786431	1	US-10-027-983-11
C	29 169	7.4	161652	1	US-09-816-095-3
C	30 149.3	6.6	70000	1	US-09-751-389-3
C	31 145.4	6.4	64467	1	US-09-497-855A-40
32	142.8	6.3	1540	1	US-09-851-896-3
C	33 136.8	6.0	9573	1	US-09-803-671B-3
					US-09-220-132-168

34 125.5 5.5 50000 1 US-09-146-053-4 Sequence 4, Appli
35 117.1 5.1 3023 1 US-09-525-160B-9 Sequence 9, Appli
C 36 117.1 5.1 3213 1 US-09-525-160B-4 Sequence 4, Appli
C 37 116 5.1 116592 1 US-09-818-512-3 Sequence 3, Appli
38 115.3 5.1 38564 1 US-09-734-673-3 Sequence 3, Appli
39 108.3 4.8 8224 1 US-09-010-398-14 Sequence 14, Appli
40 108.3 4.8 8224 1 US-09-366-260-14 Sequence 14, Appli
41 106.2 4.7 468 1 US-09-702-705-604 Sequence 604, App
42 106.2 4.7 468 1 US-09-736-457-604 Sequence 604, App
43 106.2 4.7 468 1 US-09-614-124B-604 Sequence 604, App
44 106.2 4.7 468 1 US-09-671-325-604 Sequence 604, App
45 106.2 4.7 468 1 US-09-589-184-604 Sequence 604, App

ALIGNMENTS

RESULT 1
US-08-781-891-207
; Sequence 207, Application US/08781891
; Patent No. 6090620
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; APPLICANT: Yu, Chang-En
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.
; APPLICANT: Schellenberg, Gerald D.
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
; TITLE OF INVENTION: WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,891
FILING DATE: 27-DEC-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: No. 6090620tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 240052.419
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 207:
SEQUENCE CHARACTERISTICS:
LENGTH: 29604 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-781-891-207

Query Match 25.8%; Score 588.3; DB 1; Length 29604;
Best Local Similarity 52.9%; Pred. No. 0.24;
Matches 1170; Conservative 0; Mismatches 567; Indels 123; Gaps 18;

QY 292 TCATTTCCTGATTTCTATCTGGCTCATTTTAATCATGATGATGTTGGTTTCC 351
Db 23722 TTCTTCTTATTCCTCTTCCACCAAGATCATTTGAGAGATGTTTTCAGTTTCC 23781
QY 352 ATAAGTTTGAATTTCTCTGTTTCTGTTGTTGTTGTTTATCTAGATTTAAGCTGT 411
Db 23782 ACCTGAATGTTGCTTTCATTATTATGTTGTTTGAAGATC---AGCCTTAGCCAT 23838

QY	412	GGTGGT	CAGATAGGACATAGAGTATATATTTCAATTGCTCTTTTAICTGTCGAGACTTGCTTT	471
Db	23839	GGTGGT	CTGATAGGATACATGGGACAAATTTTCAATATTTTGTGTATCTATTTGAGGCGCTGT	23898
QY	472	TGTTTTGA	ATAATGATTAACAAATTTTGGGAGA--GTUATCATAGGCTCTGACAGAAGGTAC	529
Db	23899	TGTACCA	ATTATATGGTCAATTTTGGAGAAGGTCCCGTAGGTGCTGAGAAGAAGGTAT	23958
QY	530	AGTC	-TTTGTGTTTTGGTGAATAAGTCTGTATAATACTCT-AGTCCACTTTGGTTTTATGA	587
Db	23959	ATCC	TTTTGTTAGGATAAAATGTTCTGTAGATATCTGTCAAGTCCATTTGTTTTCATAA	24018
QY	588	CATCAG	TAGTCCACATTTCTCTGTGTTGTTTTTGTGAGATGACCTAACTGTTGGA	647
Db	24019	CTTCTG	TATGTTCACTGTGCTCCCTGTTTAG-TTTCTGTTTCCAGATCTGTCTCTTTGAA	24077
QY	648	GAGAA	TGGGATTTGAAGTAGCCCACTATCTGTGTGAGGT-CAATATGTGATTTTAGC	706
Db	24078	GAAAG	TGGTGTGAAGTCTCCCACTATTATTGTGTGAGGTGCAATGATGCTTTGAGC	24137
QY	707	TGTAG	CTGTGTTTTATGAACCTTGGTGACATGTGTTGTGTGCAATAGACATTAAGA	766
Db	24138	TTTACT	AAAGTGTCTCTAATGAATGTGGCTGCCCTTGCAATTTGGTGGCTAGATATTAGA	24197
QY	767	ATTGCA	ATCTCTTGTGGGATTTT-CCTTTGATGCCATGTAGTATCTTTTCCCAATCT	825
Db	24198	ATTGAG	TGTTCTTGGAGATTTTACCTTTGATGATGAAGTGTCCCTCCCTGTCT	24257
QY	826	CATCT	CTTAGTTTTGGGTTTTAAGTCTA-TTAGTCAGATATAAAATGACTGTATCGGCT	884
Db	24258	TTTTTG	ATAACTTTGGTGTGGAGTCAATTTTATCCGACTACTAAATGGCTACTCCAGCT	24317
QY	885	TGCTTC	TTAGGCCATTTCTCTTAGAATA-TCTTTTTCCATCTCTTTTACTTAAGTGTAT	943
Db	24318	TGTTCT	CTCAGTCCATTTGCTTTGGAAAATGTGTTTCCAGCCCTTTTACTCTGAGGTAGT	24377
QY	944	CTATCC	AT---GATAGTTGTCTTTTTTGGATGCAGCAGTAGGATGGATCTGTTTTCAT	1000
Db	24378	CTGTCT	TTTTTCCCTGAGATGGGTTTCCCTGAAGCAGCAGAAATGTTGGGCTCTGTTGTGT	24437
QY	1001	ATCCAT	CTGTGTACCAGTATCTTTTTTCTAGAGAAATTAAGATCATTCAGTCAATTGAT	1060
Db	24438	AGCCAG	CTGTGTAGTCTATGTCTTTTTATGGGAAATTAG-----TCCATTTGAT	24489
QY	1061	TGAGAA	TATCAATGACAGTGTGTTGTGATTTCTGTGTTATCTGCACTGTGCAAGTGTGT	1120
Db	24490	TAAGAG	ATATAAGGAAAAGTAATGTTGTTCCCTTTTATTTTTGTTGTAGAGTTGGCA	24549
QY	1121	GTGTGT	GT	1180
Db	24550	TTCTGT	CTGTGTGGCTTCTCTTTTGGTTTG-----	24582
QY	1181	GTGTCT	GTGTGTGTGTCTCTCCCTCTTTTGATTTTTGGGCTGGAAATAT	1240
Db	24583	-----	-----TTGAATGAT	24591
QY	1241	TTATTA	TATCATATTTCTTTTGAATGGGGTAAACATCTTTAGATTTGAAGTTTTTCTCCTAGC	1300
Db	24592	TACTTT	CTTGGTTGTTCTAGGGCGTGAATTCGGTTCTTGTATTTGCTTTCTTTCTGTTAT	24651
QY	1301	CTTCTTT	-AGTCTGCATTTGAAGATAGATATCTTTACATCTCATTTTATCTTAGAAT	1358
Db	24652	ATCC	TTTGAAGGGCTGGATCTGGAAAGATATTGTGTGAATTTGTTTTGTCTGGAT	24711
QY	1359	GTCTTT	CTTTCTTCTTCCAACTATTGTGACAGAAAGTTTTTTTAAAGTCAGTGTCTGGCCTGA	1418
Db	24712	ACTTTG	TTTCTCCATCTATGGTAATTTGAGACTTTTGGGCTGGTGTATGATGCTGGGCTGG	24771
QY	1419	CATCTG	TAGTCTTGGAGTCTGTAGCACATCTGTGCGGGCCCTCTTTACATTTTGAGTT	1478
Db	24772	CATTTG	TGTTCTTTAGTTTCTGTATATAACATCTGCCAGGCTCTTCTGGCTTTTCAATGTC	24831
QY	1479	TCTATT	TGGAAAAGTCAGGTGTAAATCTCAATCATCTGCTTTTATATGTAAATTTGGTCTTT	1538

RESULT 2

RESULT 2
US-09-618-166-207

US-09-018-108-207
; Sequence 207, Application US/09618166

; Patent No. 6583112

GENERAL INFORMATION:

APPLICANT: Fu, Ying-Hui

Yu, Chang-En

Oshima, Junko

Mulligan, John T.

----- Schellenberg, Gerald D. -----

;;
TITLE OF INVENTION: GENE AND GENE PRO-

WERN

; NUMBER OF SEQUENCES: 20;
; CORRESPONDENCE ADDRESS:

; CORRESPONDENCE ADDRESS: seed Intellectual Property Law Group

ADDRESSEE: seed Intellectual Property
STREET: 701 Fifth Avenue, Suite 6300

STREET: 701 FLICK
CITY: Seattle

CITY: Seattle
STATE: Washington

STATE: WASHINGTON
COUNTRY: USA

COUNTRY: USA
ZIP: 98104-7092

COMPUTER READABLE FORM:

```

/ ;
MEDIUM TYPE: Floppy disk

```

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

```

;
;
CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/09/618,166

FILING DATE: 17-Jul-2000
CLASSIFICATION: ~~Unknown~~

US-08-781-891-209/c
; Sequence 209, Application US/08781891
; Patent No. 6090620
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; APPLICANT: Yu, Chang-En
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.
; APPLICANT: Schellenberg, Gerald D.
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
; TITLE OF INVENTION: WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESS: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781.891
; FILING DATE: 27-DEC-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6090620tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 240052.419
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 209:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 51259 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-781-891-209

Query Match 24.8%; Score 565.399; DB 1; Length 51259;
Best Local Similarity 62.1%; Pred. No. 0.18;
Matches 1155; Conservative 0; Mismatches 576; Indels 128; Gaps 18;

QY	292	TCTATTCTTGATTTCTATCTTGGCTCATTTTAACTCAGTAGTGAGTTGTTGGTTTC	351
DB	10412	TCTTCATTTTCAATTTCTGCTGACCAAGGTATCATTCAGTAGAGCGCTGTTCA	10353
QY	352	ATAAGTTGTAAAGTTTCTGTTGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTG	411
DB	10352	ATATCTATATGTTGTTTCCGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	10293
QY	412	GGTGTCTAGATAGGACATAGAGTATTATTCATTTGTTCTTTTATCTGTCGAGCT	471
DB	10292	GGTCTGATGATAGGTCATGGGATGATTTCAATCATCTTGAATCTGTAGAAGCT	10233
QY	472	TGTTTTGAAATGATGATTTCAATTTTGGAGA--GTTTCATAGGGTGTGACAGAA	529
DB	10232	TGTGACCACTATATGTCAGTTTGGAGAGGTTTCCATGAGGTGCTGAGAGAGG	10173
QY	530	AGTCTTTGTTGTTTTCGTTGAATAGTCTGTAATATCT--CTAGGTCCACTG	588
DB	10172	ATTTATGCTTTTGGATGACATGTTCTAATAATCTGTTAGATCCATTTGTTCA	10113
QY	589	ATCAGTTAGTCCAGCATTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTG	648
DB	10112	ATCTGTAGTTTCACTGTGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	10053
QY	649	AGAATGGGGTATTGAAGTAGCCCACTATCTGTGTGAGGT--CAATATGTGAT	707

Db	10052	AGAGTGGGTGCTGAAATCTCCCACTATTATTTGTAOCGGTGTGATGTGCTTTGAGAT	9993
QY	708	GTAGCTGTGCTTTGTTTATGAACCTGGGTGACATTTGTTTGGTGCATAGACATTAAGAA	767
Db	9992	TTAGTAAAGTTTCTTTTATGATGTTGGGTGCCCTTGCATTTGTAGCATACATGTT	9933
QY	768	TTGCAATGTCCTCTTGGTGAT--TTTCTTTGATGCCCTATGATGATTTCTTCCCA	826
Db	9932	TTGAGAGTTTCATCTTGGCAGATGTTTCTTCTTGCACAGTATGAAGTGTCTCTT	9873
QY	827	ATCTGCTTAGTTTCTGGGTTT-----AAGTCATTAGTCAGATATTAATAAGC	880
Db	9872	TTTTTTTTTTTGAATACTTGGTTGAGATTTGAATTTATCCATATAGATGGCTAC	9813
QY	881	GGCTTGCCTTCTTAGG-----GCCATTTGCTTAGAATA--TCATTTCCATCT	934
Db	9812	AGCTTGTCTTCTTGGGAAACAACCAITTTGCTTGGAAATTTGTTTCCAACTTGA	9753
QY	935	AGGTGATGTCATCCATG---GTAGGTTGTCCTTTTGGATGCAGCAGTAGGATG	991
Db	9752	AGGTAGTGTCTGCTTGTGCACTGAGGTGCATTTCTCTGTATGCAGAAAATGCT	9693
QY	992	TGTTTTCATATCCATTTCTGTTTACCAGATATCTTTTCTAGAGAAATTAAGAT	1051
Db	9692	TGCTTACACATCCAGTCTGTAGTCTAATGCTCTTTTGGAGAAATGA-----	9642
QY	1052	CATTGATGTTGAGAAATTAATCAATGACGAGTGTGTTGGATTTCTTCTTA--	1110
Db	9641	CATTGATGTTAAGAGATATTAAGGAAAAGTGAATTTACTTCTCTGTTATTTT	9582
QY	1111	TGAAGT	1170
Db	9581	TTAGAGTGGAAATATGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	9537
QY	1171	GT	1230
Db	9536	-----	9537
QY	1231	CTGGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT	1290
Db	9536	-TTGAAAGATTTACTTTCTTGTCTTCTTAGGGTGTAGTTTCCCTCTTGTGTGT	9478
QY	1291	TTCTCCTA--GCCTTCTTTAGGTCTGCATTTGAAGATAGATATTTTACATCT	1348
Db	9477	TCCATCTATTATTCCTTTTATAGAGCTGGAATTTGTGAATGATATTTGTAATA	9418
QY	1349	ATCTTAGATGCTTTCTTTCTCCAACTATTTGACAGAAAAGTGTCTTCTAAGT	1408
Db	9417	GTATGGAATATCTTGGTTTCTCCATCTATGATTAATTCAGAGTTTGTCTGG	9358
QY	1409	TCTGGCTGACATCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTCCAGGCGCT	1468
Db	9357	CTTGGCTGGCATTTGTGTTCTTTTGGTCTGTATGACATCTGCTAGGATCTTCT	9298
QY	1469	ATTTGAGTTCTATTTGGAAGTCAAGGTGTAATTTCTAATACATCTGCTTTAT	1528
Db	9297	TTTCATAGTCTCTGTTGAGAGTCTAGTCTCTATCTGATAGGTCTGCTTTAT	9238
QY	1529	ATTTGTTCTTTTCTCCCTGCATCTTTTAAATATTTCTTTCTTTCTATACIT	1588
Db	9237	CTT-GACTTCTTTCTCCCTTACTGCTTTCTATTTCTTTCTTTTAAATGCA	9179
QY	1589	TTGATTTATGACATGTTGGGAGTTTCTTTTCCGGTCCAAATCTATTTGGTGT	1648
Db	9178	TCAATTTATGTCAGCAGGAGGA-ATTCTTTTCTTGTCCAGTCTATTGGAGT	9120
QY	1649	GCTTCTTGTACCTTGTAGAGCATCTTTTCTCAAGTTAGGAAATTTTCTTTT	1708
Db	9119	GTTTCTTGTATGTTTCAAGACATCGTTTCTTTAGTTATGGAAGTTTCTTCTA	9060
QY	1709	TTCTTGAATAATTTTCTTCTGCTTTTGACCT--GCCTTCTTCCCTTCTCTAT	1763
Db	9059	TTGTTCAAGATATTTTACTGGCCCTTTGAGTTGGGAATCTTCACTTTCTTCTA	9000

QY 1764 -----TTTGGTTTTGATAGTGTCTGTGGCTTCCGTGATGTTTTATGCTGGAAT 1814
DB 8999 ATCCTTAGGTTTCATTTTCTCATTTGTGCTCATATTTTCTGGATGTTTTGGTTAGGAGC 8940
QY 1815 ATTTTAGACTTAAATTTTCTTTGACCAAGGTATCCATTTCTTATCTTGTCTCACTG 1874
DB 8939 TTTTTCGATTTTGTATTTTCTTTGACTTGTGTGCAAGTTTTCTATGTAATCTCTGCA 8880
QY 1875 CTTGAGATCTCTCTTCTATCTTTGTTATTTCTGTGAGGCTGTCTCTGAGGTTTCT 1934
DB 8879 CTTGAGATCTCTCTTCTATCTTTGTTATTTCTGTGTTGATGCTTTGTATATATGACTGCT 8820
QY 1935 G-TTGGGTTCTAAATTTTCTTCAATTCAGATTTCTTCAAGTTTGGTTTGTATTAAT 1993
DB 8819 GATCTTTTCTAGTTTTCTGTCTCCAGCTGTGTCTCCCTTTGTGATTTCTTTTATGTT 8760
QY 1994 TCTATTTTCCAGTTTCCAGTCTCTGAAATGTTTTTACTCAATTTTCTCCCTCC---AGTATTTACA 2050
DB 8759 TCTAGTTCCATTTTAGATCTCTGGATGTTTTGTTTCAATTTCTTCACTGTTGATTTG 8700
QY 2051 TTTTCATAGGTTTCTTTAATGGAATTTATCAATTTCTTCTTCAAGGACTTTTATGAAT 2109
DB 8699 TTCTCTGTAGTTTCTTTAAAGATTTTGTGTCTTCTTCTTAAAGGCTTCTAGTTGTT 8641

RESULT 4

US-09-618-166-209/c
; Sequence 209, Application US/09618166
; Patent No. 6583112
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; Yu, Chang-En
; Oshima, Junko
; Mulligan, John T.
; Schellenberg, Gerald D.
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
; WERNER'S SYNDROME

NUMBER OF SEQUENCES: 209

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed Intellectual Property Law Group
STREET: 701 Fifth Avenue, Suite 6300

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/618,166

FILING DATE: 17-Jul-2000

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Mcmasters, David D.

REGISTRATION NUMBER: 33,963

REFERENCE/DOCKET NUMBER: 240052.419C1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 209:

SEQUENCE CHARACTERISTICS:

LENGTH: 51259 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 209:

US-09-618-166-209

Query Match

Best Local Similarity 24.8%; Score 565.399; DB 1; Length 51259;

Pred. No. 0.18;

Matches 1155; Conservative 0; Mismatches 576; Indels 128; Gaps 18;
QY 292 TCTATTTCTGATTTCTATCTTTGGCTTCATTTTAACTCAGTAGTGTGTTGGTTTCTCC 351
DB 10412 TCTTCATTTTCAITTTCTGCTTTGACCAAGGTATCATTTAGTAGAGCGCTTTCAGCTTC 10353
QY 352 ATAAATTTGTAAGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTT 411
DB 10352 ATATGATGTGTGCTTTCCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTT 10293
QY 412 GGTGGTTCAGATAGGACATAGATTTATTTCAATTTCTTTTATCTTTCTGTCGAGACTTGTCT 471
DB 10292 GGTCTGATGATAGGTTGATGAGTTCATTCATCATCTTGAATCTGTAGAGTCTGTTT 10233
QY 472 TGTGTTGAAATATGATTTCAATTTTGGAGA--GTTTCATAGGTTGCTGCAAGAGGTAC 529
DB 10232 TGTGACAGCTATATGTCAGTTTTCAGTTTTCAGAGAGTTTCCATGAGTGTGCTGAGAAGAT 10173
QY 530 AGTCTTTGTTGTTTGGTGAATAGTCTGTAATAATCT-CTAGTCCACATTTGGTTTATGAC 588
DB 10172 ATTATTTGCTTTTGGATGACATGTTCTATAAATATCTGTTAGATCCATTTGTTTCTAATC 10113
QY 589 ATCAGTTAGTCTCAGCATTTCTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTT 648
DB 10112 ATCTGTTAGTTTCACTGTGCTCTGATTTAGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTT 10053
QY 649 AGATGGGTATGAGTAGTACCCACTACTCTGCTGTGAGGT-CAATATGTTGATTTAGCT 707
DB 10052 AGATGGGTGCTGAAATCTCCACTATTTATGATCGGGTGTGATGTTGCTTTGAGAT 9993
QY 708 GTAGCTGTGCTTTGTTTATCAACTTGGGTGACATTTGTTTGGTGCATAGACATTAAGAA 767
DB 9992 TTAGTAAAGTTTCTTTTATGAATGTGGTGCCTTGCATTTGTAGCATACATGTTTCAGAA 9933
QY 768 TTGCAATGCTCTTGGTGGAT- TTTCTTTGATGCTATGATGATTTCTTCCAACTTC 826
DB 9932 TTGAGAGTTTCACTTGGCAGATTTTCTTTGACCAGTATGAAGTCTCCTCTTATCTT 9873
QY 827 ATCTGCTTAGTTTGGGTTT-----AAGTCTATTAGTCAGATTTTAAATGACTGTATC 880
DB 9872 TTTTCTTTTGTATTAATCTTGGTTGAGATTTGAATTTATTTCCATATTAGATGGTACTCTC 9813
QY 881 GGCTTCTCTTTAGG-----GCCATTTGCTTGAATA-TCCTTTCCATCTTTTACTCTA 934
DB 9812 AGCTTGTCTTTGGGAAACAACCAATTTGCTTGGAAAATTTGTTTCCAACTTGAACCTG 9753
QY 935 AGTGATGCTATCCATG---GTAGTTTCTTTTGGATGCGAGTAGGATGATCT 991
DB 9752 AGTAGTGTCTGCTTTGTCTGCTGAGTGTCTGCTGATGGCAATTTCTGTATGGCAAAATGCTGGGCTC 9693
QY 992 TGTCTTTCATATCCATTTCTGTTTACCCAGTATCTTTTCTAGAGAAATTAAGATCAITGAGT 1051
DB 9692 TGCTTACATCCAGTCTGTAGTCTATGCTCTTTTGTGAGGAATGA-----GTC 9642
QY 1052 CATTTGATTTGAGAAATTAATCAATGAGCAGTGTGTTGGAATCTTGTGTA-TCCTGCACTTG 1110
DB 9641 CATTTGATTTGAGAGATA-TAAGGAAAGATGTTGTTACTTCTGTTATTTTGTGTTG 9582
QY 1111 TGAAGT 1170
DB 9581 TTAGAGTGAATATGTTTGTGTGCTATCTTTTGGGTTG----- 9537
QY 1171 GT 1230
DB 9536 ----- 9537
QY 1231 CTGGAATTTATTTATTTTATTTTCTTGAATGTGGGTAACATCTTTAGATTGAGTTT 1290
DB 9536 -TTGAAAGATTACTTTCTTGTCTTCTTAGGGGTAGTTTCCCTCTTGTGTGTGTGTGT 9478
QY 1291 TTCTCTCA--GCCTTCTTTAGTGTGCAATTTGAAGATAGATATCTTTACATCTGATTTT 1348
DB 9477 TCCATCTATTATCTTTTGTAGAGCTGGAATTTGTGGAATCATATTGTGTAATTTGTTTTT 9418

Qy	941	TGCTATACGAG---GATAGGTGTCTTTTTTCGATGCGAGCAGTAGGATGGATCTCTGTTTT	997
Db	26759	TGTCCTGCCCTTGTCACTAGGTGCAATTCCTCTATGACGACAAATGCTGGGTCTCTGTTTA	26918
Qy	998	CATATCCATTCTGTTACCCAGATCTTTTTCTAGAGAAATTAAGATCAATTGAGTCATTGA	1057
Db	26819	CACACCAGTCTGTAGTCTATGTCTTTTTTTCAGGAATTGA-----GTCCATTGA	26869
Qy	1058	TGTTGAGAAATTAACAAGCAGAGTGTGTGGGATCTTGTTATCTTGCACTGTGGAAGTG	1117
Db	26870	TGTTAAGAGATATAAGGAAAAAGTCATTGTACTTCCCTGTTATTTT-----TGTTG	26920
Qy	1118	TGNGT	1177
Db	26921	TTGTTAGAGGTGGAATATATGTTTGTGTGCTATCTCTTTTGGGTTTGTGAAGAATGTC	26980
Qy	1178	TCGTGTCTGTGTCTGTGTGTGTGTGTGTCTCTCCCTCTTTTGTGATTTTTGGCCCTGSAAT	1237
Db	26981	TTTTCTTGCTTTTCTAGGGTGTAGTTTCCCTC-----CTTGTTGTGTGTGT	27025
Qy	1238	TATTTATTATTCATATTTCTTGAATGTGGGTAAACATCTTTAGATTGAAGTTTTCCTCT	1297
Db	27026	TTTCCATCTATATCTCTTTTAGAGCTGGAAGAATTTGTGTAATTTGGTTTGTCTAAG	27085
Qy	1298	AGCCTTCTTTAGTCTGTGCATTTTGAAGATAGATATCTTTACATCTGATTTTATCTTAGAA	1357
Db	27086	AAATACCTAGACGCTTGACAGCACACCTGAACACCTAGACCTAAAAGACCAATACAC	27145
Qy	1358	TGTCCTTTCTTCTCCAACTATTGTGTGACAGAAAGTTTTTCTAAGTGCAGTAGTCTGGCCTG	1417
Db	27146	CCAAGAGGAGTAGACCTGAGATTTGGGAGTT-----TTGCCCTGGGCTG	27186
Qy	1418	ACATCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGACGGCCCTCTCAATTTTGAGT	1477
Db	27187	GCAITTTGTGTTCTCTTAGGGTCTGTATGACACTGCCTAGGATCTTTTAGCTTTCATAGT	27246
Qy	1478	TTCTATTGGAAGAGTCAGGTGTAATTTCTAATACATCTGCCCTTTATGTTAATTTGGTCTT	1537
Db	27247	TTCTGTGT-GAGAAAGCTGTGTGTAATTTCTGATAGGCTGCTTTATGTTACTTGACCT-	27304
Qy	1538	TTTTTCCCTTGCACTTTTAAATTTCTTCTTGTGTTCTATACTTTTTAGTGATTTTGATTAAT	1597
Db	27305	-TTTTCAATCTGCTTTTTTAATATCTTTCTTTGTTAGTGCATTTGGTGTGTTTGTATTAAT	27363
Qy	1598	ATGCATGTGGGGAGTTCTTTTCCGGTCCAACTATTTTGGTGTGTTTGTATGCTTCTTGT	1657
Db	27364	ATGTGACAGAGGAAATTTCTTTTCTGGTCCAGTCTATTTGGAGTTCTGGAGGCTTCTTGC	27423
Qy	1658	ACCTTGATAGGCATCTTTCTCAAGGTTAGGAATTTTTCTTTTGGTTTCTTGAAA	1717
Db	27424	ATGTTCAAGGGCATCGCTTTTTTAGGTAGGGAAGTTTCTCTCATAAATTTTGTGAAG	27483
Qy	1718	ATATTTTCCCTGTCTTTTGACCTGCCCTCTTCCCTCTCTATCTCC-----1763	
Db	27484	ATATTTACTGGCCCTTTGAGTTGGGAATCTTCACCTCTCTATACATATATCCCTAGG	27543
Qy	1764	TTTGGTTTTTGCATAGTGTCTCGGCTCTCTGAGATGTTTTATGCGCTGGAATATTTTAGAC	1823
Db	27544	TTTTGGTCTTCTCATTTGTCTCGGATTTCTCGATGTTTTTGGGTTAGGAGCTTTTTTGAT	27603
Qy	1824	TTAAACATTTTCTTTTGACCAAGGTATCCATTTCTTCTATCTTCTCTTCACTGCCTCAGATT	1883
Db	27604	TTTGTATTTTCTTTTGACTGTTGTGTCAATATTTCTATGGAATCTTCTGCACCTGAGATT	27663
Qy	1884	CTCTCTTCTATCTCTGTATTCGTG-TCAGTGAAGCTGTCTCTGAGGTTCCGTG-TTGGGT	1941
Db	27664	CTCTCTTCTATCTCTGTATTCGTGTTGGTGTGATGCTTGCAATCTCTGACTCTCTGATCTCTT	27723
Qy	1942	TCCTAATTTTTTTCATTTCCAGATTTCCCTTCAGTTGGGTTTTGTTTATTAATTTCTATTTC	2001
Db	27724	TCCTAGATTTTCTAACTCCAGGGTGTGTCTCCCTTTGTGATTTCTTTATTGTTTCTAGTTC	27783
Qy	2002	CACTTTTCAGTCTCGAAATGTTTTTACTCATTTTTTCTCCCGAGTATTAGAC--TTTTTCATA	2058

RESULT 6

US-09-618-166-209
 ; Sequence 209, Application US/09618166
 ; Patent NO. 6583112
 ; GENERAL INFORMATION:
 ; APPLICANT: Fu, Ying-Hui
 ; Yu, Chang-En
 ; Oshima, Jurko
 ; Mulligan, John T.
 ; Schellenberg, Gerald D.
 ; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
 ; WERNER'S SYNDROME
 ;
 ; NUMBER OF SEQUENCES: 209
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Seed Intellectual Property Law Group
 ; STREET: 701 Fifth Avenue, Suite 6300
 ; CITY: Seattle
 ; STATE: Washington
 ; COUNTRY: USA
 ; ZIP: 98104-7092
 ;
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/618,166
 ; FILING DATE: 17-Jul-2000
 ; CLASSIFICATION: <unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McMasters, David D.
 ; REGISTRATION NUMBER: 33,963
 ; REFERENCE/DOCKET NUMBER: 240052.419C1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (206) 622-4900
 ; TELEFAX: (206) 682-6031
 ; INFORMATION FOR SEQ ID NO: 209:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 5125 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 209:
 ;
 ; US-09-618-166-209

	Query Match	22.9%	Score 521.797	DB 1	Length 51259
	Best Local Similarity	60.7%	Pred. No. 0.3		
	Matches 1118	Conservative 0	Mismatches 637	Indels 88	Gaps 17
Qy	292	TCATTTCTTCGATTTCTATCTTGGCTCATTTTAACTCAGTAGAGTGTTGTTGTTTCC	351		
Db	26099	TCATCATTTTCATTTCTGCCCTGACCAAGTTATCATTCAGTAGAGCGCTGTTTCAGCTTCC	26158		
Qy	352	ATAAGTTGTGAAGTTTTCGTGTGTTTCTGTGTGTTGTTGTTTATCTAGATTTTAAGCTGT	411		
Db	26159	ATATGTATGTGTGCTTTCCGTTGTTTGTGTGGTATTTAAGACCAACCTTAGTCGCTGT	26218		
Qy	412	GGTGGTCAGATAGACATAGAGTATTATTCAAATGTCTTTTATCTGTCTGAGACTTGTT	471		
Db	26219	GGTCTGATGATAGGGTGATGGGATGATTTCCATCATCTTGAATCTGTAGAAGTCTGTT	26278		
Qy	472	TGTTTTGAAATATGATTTCAATTTTGAGA--GTTTCATAGGGTGCTCACAAGAGGTAC	529		
Db	26279	TGTGACCAGCTATATGGTCAGTTCTTTTGAGAAAGGTTCCATGAGGTGCTGAGAAGAGGTAT	26338		

QY 530 AGTCTTTGGTTTGGTGAATAGTCTGTAATATCT-CTAGGTCCACTTGGTTTATGAC 588
Db 26339 ATTTTGTCTTTGGATGACATGTTCTATAATAATCTGTTAGATCCATTTGGTTCATAAC 26398
QY 589 ATCAATAGTCCAGCATTTCTCTGTTTGGTTTGGTTTGGTGAAGTGAAGTGAAGTGAAG 648
Db 26399 ATCTGTAGTTTCACTGTGCTCTGCTGTAGTTTCTGTTTCCGTTGATCTGTCCTTCCATG 26458
QY 649 AGAATGGGATTTGAAGTAGCCCACTATCTGTGTGAGGTCATATGTGTGATTTAGCTG 708
Db 26459 AGAGTGGGCTGTAATATCTCCCACTATATTTGATCAGGTATGATGTGCTTTGAGAT 26518
QY 709 TAGCTGTGCTGTTTATGAACCTGGGTGACATTTGTTTGGTGCATAGACATTAAGAAT 768
Db 26519 TTAGTAAAGTTTATGATGATGGGTGCCCTTGCAATTTGGAGCATACATGTTTCAAGAT 26578
QY 769 TGCAATGCTCTGTTGGTGGAT-TTTCCTTTGATGCTATGTATGTTTCCCAATCTCA 827
Db 26579 TGAGAGTTTCATCTTGGCAGATGTTTCCCTTTGACCAATATGAAGTGTCTCTCTTATCTTT 26638
QY 828 TCTGCTTAGTTTGGGTTTAAAGTCT-ATTAGTCAATATTAATGACTGTATCGGCTTG 886
Db 26639 TTTTGTAACTTGGTTCAGAGTTGAATTTATCCATATTAGATGGCTACTCCAGCTTG 26698
QY 887 CTCTTAGG- - -GCCAATTTGCTTAGAATA-TCTTTTCCATCTCTTTTACTCTAAGGTGA 940
Db 26699 TTTCTTGGGAAACACCACTTTGCTTGGAAATTTGTTTCCAACTTGAACCTCTGAGGTAG 26758
QY 941 TGCTATCATG- - -GTAGGTTGCTTTTGGATGAGCAGTAGGATGGATCTTGTGTTT 997
Db 26759 TGCTGCTTTGCTACTGAGGTGCAATTTCTCTGATGACGAAATGCTGGTCTCTGTTTA 26818
QY 998 CATATCCATTTCTGTTACCAGTATCTTTTCTAGAGAAATTAAGATCAATGAGTCATGA 1057
Db 26819 CACACCCAGTCTGTAGTCTATGCTTTTGTAGGAATTC- - -GTCCATTA 26869
QY 1058 TGTTGAGATTAATCAATGAGCAGTGTGTTGGATCTGTTATCTTGCACTTGTGAAGT 1117
Db 26870 TGTTAAGAGATTAAGAAAGTGAATGTTACTTCTGTTTATTTT- - -TCTTG 26920
QY 1118 TGTTGCTGTGTTGCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1177
Db 26921 TTGTTAGAGTGGATTAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 26980
QY 1178 TCTGTGCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1237
Db 26981 TTTCTTGTCTTTTCTAGGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 27025
QY 1238 TATTTATTAATCATATTTCTTGAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1297
Db 27026 TTTCCATCTATTTCTTTTAGCTGGAAGATATGTTGTTGTTGTTGTTGTTGTTGTTGTTG 27085
QY 1298 AGCCTCTTTAGTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1357
Db 27086 AAATACCTTAGCAGTGTGACGACACACCTGAACTCTAGAACTAAAGAAAGCAATACAC 27145
QY 1358 TGCTTTCTTCTCACTATTTGACAGAAAGTTTCTTAAGTGCAGTAGTCTGGCCTG 1417
Db 27146 CCAAGAGAGTAGCTGAGATGGAGTT- - -TTGCCCTGGGCTG 27186
QY 1418 ACATCTGTAGTCTCTTGGAGTCTGTAGCACAATCTGTGAGGCGCTCTTATCATTTTGA 1477
Db 27187 GCAATTTGTTCTCTTAGGCTGTATGACATCTGCTTAGATCTTTTAGCTTTTCAATG 27246
QY 1478 TTCTATTTGAAAGTCAAGTGAATTTCTTAATCATCTGCTTTTATATGTTTAAATGTTCT 1537
Db 27247 TTCTGGT-GAGAAGTCTGGTGAATTTCTGATAGGCGCTGCTTTATATGTTTACTTGA 27304
QY 1538 TTTTCCCTTGAATCTTTTAAATTTCTTTTCTTGTCTATATCTTTTAGTGAATTTGATTA 1597
Db 27305 -TTTCCATTTGCTTTTAAATTTCTTTTCTTTGTTAGTGAATTTGTTGTTGTTGTTG 27363
QY 1598 ATGCACTGGGAGTTTCTTTTCCGGTCCAATCTATTTGGTGTGTTGTTGTTGTTGTTGTTG 1657

Db 27364 ATGTGACAGGAGAAATTTCTTTTCTGTCAGTCTATTTGGAGTTCTGGAGCTTCTTGC 27423
QY 1658 ACCTTGATAGGATCTCTTTCTCAAGGTTAGAAATTTTCTTTTGGTTTCTTGA 1717
Db 27424 ATGTTTCATGGGATCGCTTTTCTTAGGTTAGGAAAGTTTCTTCTATAATTTTGTGA 27483
QY 1718 ATATTTTCTCTGTTTGGACCTGCTCTTCTTCCCTTCTCTATTC- - - 1763
Db 27484 ATATTTACTGGCCCTTGGATTGGGAATCTTCACTCTCTTCTATACATATATCCTTAG 27543
QY 1764 TTGGTTTGGATAGTGTCTCTGCTCTCTGCTGATTTTATGCTGATTTATTTAGAC 1823
Db 27544 TTGGTCTTCTCATTTGTCTCTGATTTCTGATTTTGGGTTAGGAGCTTTTTCAT 27603
QY 1824 TTAACATTTTCTTGGACCAAGGATCCATTTCTCTATCTTCTCTCTCTGCTGCTGAGATT 1883
Db 27604 TTGTTATTTCTTGGATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 27663
QY 1884 CTCTCTTCTATCTTGTATTTCTG-TCAGTGAGGCTTGTCTCTGAGGTTCTG-TTGGGT 1941
Db 27664 CTCTCTTCTATCTTGTATTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 27723
QY 1942 TCTTAATTTTCTTCAATTTCCAGATTTCTCTGATTTTGGGTTTGGTTTATTAATTTCTATT 2001
Db 27724 TCTAGATTTTCTAACTCCAGGTTGTCTCTCTTGTGATTTCTTATTTGTTTCTAGTTC 27783
QY 2002 CACTTTCAGTCTCTGAAATGTTTACTCATTTTCTCTCCAGTATTACA- - -TTTTTATA 2058
Db 27784 CATTTTAGACTCTGATGTTTGTTCATTTCTTCTGCTGTTTAAAGTGTGTTTCTGG 27843
QY 2059 GGTTCCTTTAATGGATTTATTCATTTCTCTCTTCAAGGACCTTT 2101
Db 27844 TAATCTGTAGGAATTTTGTGTTTCTCTCTTAAAGGCTTCT 27886

RESULT 7
US-08-781-891-208
; Sequence 208, Application US/08781891
; Patent No. 6090620
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; APPLICANT: Yu, Chang-En
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.
; APPLICANT: Schellenberg, Gerald D.
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
; TITLE OF INVENTION: WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,891
; FILING DATE: 27-DEC-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6090620tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 240052.4-19
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 208:

[illegible]


```
QY 1423 TGTATCTCTCTGGAGCTCTGTAGACACATCTGTGACGGCCTCTTACATTTTGAAGTTCTTA 1482
Db 111231 TCTTTTCTTTAAGAAATGTTGAATATGGCCCTCACTCTCTCTGACTGTGAGAGTTCTTG 111290
QY 1483 TTGGAAGAGTCAGGTGAATCTTAATACATCTGCTTTATATGTTAATGTCCTTTTTC 1542
Db 111291 CTGCGAAA-TGCGCTGTAGTCTGAAGGCTTCCCTTTGT-GGGTAAACCCGACCTTCTC 111348
QY 1543 CTTGCAATCTTTTAATATCTTCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1602
Db 111349 TCTGCTGCCCTTAACATTTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 111395
QY 1603 CTGTGGGAGTTCTTTCCGGTCCAACTATTTGGGTTTGTATGCTTCTTCTTCTTCTTCTT 1662
Db 111396 TCTTGGAGTGTCTTCTCGAGGAGTATCTTGTGGGCTTCTCTGTATTTCTCTGAATCTG 111455
QY 1663 GATAGGCATCTTCTCTCAAGTGTAGGAAATTTCTTTTGTGTTCTTCTTGAATATT 1722
Db 111456 AATGTGGCGTCTGTAGATTTGGGAAATTTCTCTGGAATATATCTCTCGAGTGT 111515
QY 1723 TTCC-----CTGCTTTTGACCTGCCTTCTTCCCTTCTCTTCTTCTTCTTCTTCTTCTT 1768
Db 111516 TTCCAACTTGTTCATTTCCCTCTCACTTTTCAGGTACACCAATCAGACGTAGATTGG 111575
QY 1769 TTTTGTGATAGTCTCTGGCTTCTGATGTTTATGCTGGATTTTATGCTGGATTTTATGACTTAAC 1828
Db 111576 TCTTTTACATAGTCCCAATTTCTTGGAGCTTTGTTCAATCTTTTATCTTTTCTTCTTCTT 111635
QY 1829 ATTTCTTTGACCAAGTATCATTCTTCTTATCTTGTCTTCACTGCTGAGATCTCTC 1888
Db 111636 TCTAACTCTCTTTTCACTTCAATTTCAATTTGATCTTCCATCACTGATCTCTTTC 111695
QY 1889 TTCTATCTTGTATCTGTGAGTCTGAGGCTTGTCTCTGAG 1928
Db 111696 TTCCAGTTGATGATCTGGCTACTGAGGCTTGTGCAATTAG 111735
```

RESULT 10

```
US-09-497-855A-44
; Sequence 44, Application US/09497855A
; Patent No. 6605432
; GENERAL INFORMATION:
; APPLICANT: Huang, Tim
; TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR DETECTING DNA METHYLATION
; FILE REFERENCE: UM01523
; CURRENT APPLICATION NUMBER: US/09/497,855A
; CURRENT FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/120,592
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: 60/118,760
; PRIOR FILING DATE: 1999-02-05
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 44
; LENGTH: 193303
; TYPE: DNA
; ORGANISM: Homo sapiens;
US-09-497-855A-44
```

```
Query Match 11.1%; Score 253.097; DB 1; Length 193303;
Best Local Similarity 51.9%; Fred. No. 1.7;
Matches 861; Conservative 0; Mismatches 684; Indels 115; Gaps 14;
```

```
QY 292 TCTATTTCTTCAATTTCTGCTCATTTTAACTCAGTAGTGAGTGTGTTGGTTTCC 351
Db 110168 TTTCTGCTTCAATTTCAATTAAGTACCCAGTAGTCAATGAGGAGGAGTGTGTTCAAGTTCC 110227
QY 352 ATAAGTTTGAAGTTTCTGTTTCTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 411
Db 110228 ATGAGTTGAGTGTGTTGTTGAGT-...TTCTTAATCTCAGTTCTAGTTTGAATGCAC 110283
QY 412 GGTGTGATAGGACATAGAGTATTTTCAATTTGCTTTTATCTGTCGAGACTTCTT 471
```

```
Db 110284 TGTGCTGTGAGACACAGTTTGTGTTATTAATTTCTGGTCTTTCACATTTGCTGAGGAGTCTT 110343
QY 472 TGTGTTTGAATAATGTTATTCATTTTTCGA--GAGTTTTCATAGGCTGTGACAAAGGTAC 529
Db 110344 TACTTCCAACTAATGCTGCTCAATTTTTCGAATAGGTGTGTTGCTGCTGAAAGAAAGTAT 110403
QY 530 AGTCT-TTGTGTTTTCGGTGAATAGTCTGTAAATATCTCTAGGTCCACTTGGTTTATGAC 588
Db 110404 ATTCTGTGATTTGGGTGGAGAGTTCTGTAGATGTCTATTAGGTCTGCTGTGTCAGAG 110463
QY 589 ATCAGTTAGTCCAGCATTTCTCTGTTTCTGTTTGTGTTTGTGATGACCTAACTGTTGGAG 648
Db 110464 CTGAGTTCAATTCCTGGGTGCTCTTGTAGCTTTCTGCTTGTGTATCTGTCTAAATGTTG 110523
QY 649 AGAATGGGTATTAAGTAGTCCCACTATCTGTGTGTGAG-GTCAATATGTGATTTTAGCT 707
Db 110524 ACAGTGGGTGTAAAGTCTCCCAATTTTATTTGTGTGGAGICTAAGTCTCTTTGTAGGT 110583
QY 708 GTAGCTGTGCTTGTGTTTATGAATCTGGGTGACATTTGTGTTTGGTCATAGACATTAAGAA 767
Db 110584 CACTCAGGACTTGTCTTTATGAATCTGGGTGCTCTGTATTTGATGTCATATATATTTAGGA 110643
QY 768 TTGCAATGCTCTCTGCTGGAATTTCTCTTGTATGCTCTATGTAGTATTTCTTCCAACTCTCA 827
Db 110644 TAGTTAGCTCTTCTCTGTTGAATGTATCCCTTACCATTATGTAATGSCCTTTTGTCTCT 110703
QY 828 TCTGCTTAGTTTGGGTTTAAGTCT-ATTAGTCAGATATATAAAAGCACTGTATCGGCTTG 886
Db 110704 TTTGATCTTTGTTGGTTTAAAGTCTCTTTTATCAGAGACTAGGATTTGCAACCCCTTCCCT 110763
QY 887 CTTCTTTAGGGCCATTTGCTTAGAATATCTT-TTCCATCTCTTTTACTCTAAGGTGATGCT 945
Db 110764 TTTTGTGTTTCCATTTCTTGTGATCTTCTTCCATCTCTTTTATTTGAGCCTATGTGT 110823
QY 946 ATCCAT---GGTAGGTTGTCTTTTGTGATGAGCAGTAGTAGGATGATCTTTGTTTTCATAT 1002
Db 110824 GTCTCTGAGGTGGGATGGTTTCTCTGAATACAGCACACTGATGGTCTTGTACTCTTAT 110883
QY 1003 CCATCTGTATCCAGTATCTTTTCTAGAGAAATTAAGATCATTAGTCATGTGTTG 1062
Db 110884 CCAATTTGCCAGTCTGTGTCTTTTAAATGGAGCAATTTAGCCCAAT- 110928
QY 1063 AGAATATCAATGAGCAGTGTGTTGTGGATCTTGTGATCTTGCACTTTGGAAGTGTGTCT 1122
Db 110929 -----TACATTTAAGTTGATATTTGTTATGTGTGAAAT 110960
QY 1123 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1182
Db 110961 TTGATCCTGTCTGTATG-----ATGTTAGTCTGCTCATTTT 110995
QY 1183 GTCTGTGTCTTGTGTGTGTGTGTCTCCCTCTTCTCCCTTTTGGCCTGGAATATTTT 1242
Db 110996 GCTTGTAGTTGATGAGTGTCTTCTAGCCTCGATGGTCTTACAAATTTGCAATGTTTTT 111055
QY 1243 ATTAATCATATTTTCTTGAATGTGGGTAAACATCTTTAGATTTGAAGTTTTCCTTAGCT 1302
Db 111056 TACAGTGGGTGTTGTTGTTCTTCTTCCATGTTTAG-----TGCTTCTTTCAGAAGCT 111110
QY 1303 TCTTTAGGCTGCAATTTGAAGATAGATATCTTTTACATCTGATTTTATCTTGAATGTCT 1362
Db 111111 CTTTTAGGCGAGGTCTGGGTGACAAAATCTCTCAGCAATTTGCTGTCTGTAAGATATT 111170
QY 1363 TTCTTTCTCCAACTATTGTGACAGAAAGTTTCTTAAAGTGCAGTAGTCTGGGCTGACATC 1422
Db 111171 TTATTTCTCTTCACTTATGAAGCTTAGTTTGGCTGGATGAAATCTTGGGTGAAAT 111230
QY 1423 TGTAGTCTCTTGGAGTCTGTAGACACATCTGTGCGAGGCTTCTTTCATTTTGTAGTTTCTA 1482
Db 111231 TCTTTTCTTTAAAGATTTGAATATTGGCCCTCACTCTCTCTGACTTGTAGAGTTTCTG 111290
QY 1483 TTGGAAGCTCAGGTCTAATCTTAATACATCTGCTTTTATATGTTAATTTGTTCTTTTTC 1542
Db 111291 CTGCGAAA-TCCGCTGTAGTCTGATGGGCTTCCCTTTGT-GGGTAAACCCGACCTTCTC 111348
```

QY	1543	CCCTGCATCTTTTAATATCTCTCTTGTCTATACCTTTAGTAGATTTGATTATTATGCA	1602
Db	111349	CTCTGCTGCCTTAACATTTTTTCCTTCATTCAA-----GACAATTATGTG	111395
QY	1603	CTGTGGGAGTTCTCTTTTCGGTCCAAATCATATTGGTGTTTTGTATGCTCTGTGTACCTT	1662
Db	111396	TCTTGGAGTTGCTCTTCTCGAGGAGTATCTTTGTGGGCTCTCTGTATTATTCCTGAATCTG	111455
QY	1663	GATAGGCATCTCTTTCTCAAGGTTAGAAATTTTTCTTTTTTGGTTTTCTTGAATATATT	1722
Db	111456	AATGTTGGTGCTCTGTAGATTGGGGAATTTCTCTGGATAATAATCTCTCAGAGTGTT	111515
QY	1723	TTCC-----CTGCTTTTGACCTGCCTTCTTCCCTCTCTATTTCTCTGG	1768
Db	111516	TTCCAACTTGTTTCATTTCTCCCTGTCACTTTCAGGTACACCAATCAGACGTAGATTGG	111575
QY	1769	TTTTTGCATAGTGTCTGTGGTCTCTGGATGTTTTATGCTGGATTATTTTAGACTTAAC	1828
Db	111576	TCTTTTCACTAGTCCCATATTTCTGGAGCTTTGTCAATTTCTTTTATTCTTTTTC	111635
QY	1829	ATTTTCTTTGACCAAGGTATCCATTTCTCTATCTTCTTCACTGCCTGAGATTCTCTC	1888
Db	111636	TCTAAACTCTCTTTTCACTTCAATTTCAATTCATTGATCTTCCATCACTGACTCTTTC	111695
QY	1889	TTCTATCTCTGTATTTCTGTCAGTGAGCTTGCTCTGAG	1928
Db	111696	TTTCCAGTTGATCGATTCTGGTACTGAGCTTTGTGCATTAG	111735

RESULT 11

US-10-027-983-11/c
; Sequence 11, Application US/10027983

```

; Patent No. 6617162
;
; GENERAL INFORMATION:
;
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Mark P. Roach
;
; TITLE OF INVENTION: ANTISENSE MODULATION OF ESTROGEN RECEPTOR ALPHA EXPRESSION
;
; FILE REFERENCE: RTS-0340
; CURRENT APPLICATION NUMBER: US/10/027,983
; CURRENT FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 98
;

```


Db 318461 ATGTAGTTGACGGGTTTGTAGTGTAGT---TTCTTAATCCCTGAGTTCCTAGTTTGAATGCAC 318405
Qy 412 GGTGGTCAGATPAGACATAGAGTATTATTTCAATTTGCTTTTATCTCTCGAGACTTGCTT 471
Db 318404 TGTGGTCGAGAGATAGTTTGTATAAATTTCTGTTCTTTTACATTTGCTGAGGAGTGT 318345
Qy 472 TGTTTGAAATATGATTTCAATTTTGA--GAGTTTTCATAGGCTGCTGACAAGAGTAC 529
Db 318344 TACTTCCAACATATGTGTCATTTTGGATTAAAGTCCGGCGGTGCTGAGAAAGATGAT 318285
Qy 530 AGTCT-TTGTGTTTGGTGAATAGTCTGTAATAATCTCTAGGTCACCTTGGTTTATGAC 588
Db 318284 AATCTGTTATTTGGGTGGAGAGTCTGTAGATGTCTATTAGTCTGCTTGGTGCAGAG 318225
Qy 589 ATCAGTTAGTCCAGCAATTTCTGTTTGGTTTGGTTTGGTGGAGTCACTGTTGGAG 648
Db 318224 CTGATTCAGTTCCTGGATATCCCTGTTAAATTTTCTGCTCGTTGATCTGCTAAATTTG 318165
Qy 649 AGAATGGGTATTCAAGTAGCCCACTATCTGTGTGTGAG-GTCAATATGTGATTTAGCT 707
Db 318164 ACAGTGGAGTGTAAAGTCTCCCAATTATTATGTGTGGAGTCTAAGTCTCTTTGTTAGT 318105
Qy 708 TAGCTGTGCTGTTTATGAACCTTGGGTGACATTTGTTTGGTGGCATAGACATTAAGAA 767
Db 318104 CACTAAGGACTTGTCTATGAATCTGGGTGCTCCTGATTGGGTGCATATATTTAGGA 318045
Qy 768 TTGCAATGCTCTCTGTTGGATT-TCTTTTGAATGCTATGTAGTATTTCTCCCAATCTC 826
Db 318044 TAGTTAGTCTCTTGTGTTGAATGATCCCTTTACTATTATGTAATGGCCCTCTTTGTCTC 317985
Qy 827 ATCTGCTTAGTTTGGGTTTAAAGTCT-ATTAGTCAAGATATTAATAAGTGTATCGGCTT 885
Db 317984 TTTTGAATCTTTGTTTAAAGTCTGTTTAAACAGAGACTAGGATTGCAACCCCTGCTT 317925
Qy 886 GCTTCTTAGGCCAATTTGCTTAGAATATCTT-TTCCATCTCTTTTACTCTAAGGTGATGTC 944
Db 317924 TTTTGTGTTTCCCAATTTGCTGGTAGATTTTCCFCCATCCCTTTATTTTGGAGCCTATGTG 317865
Qy 945 TATCCATG--GTAGTGTGCTTTTGGATGACAGAGTAGGATGGATCTTGTTTTCATA 1001
Db 317864 TGTCTGTGACGTGAGATGGGTTTCTGTAATACAGACGCTGATGGGCTTGACTCTTTA 317805
Qy 1002 TCCATCTGTTTACCAGTATCTTTTCTAGAGAAATTAAGATCAATGAGTCAATGATGTT 1061
Db 317804 TCCAATTTGCCAGTCTGTCTTTTAAATGAAGCAATTTAGCCCAAT----- 317759
Qy 1062 GAGAAATATCAATGACGAGTGTGTTGTGATCTTGTGATCTTGTATCTTGCACTGTGAGTGTG 1121
Db 317758 -----TACATTTAAGGTTAATATTTGATGTGTGAA 317728
Qy 1122 TG 1181
Db 317727 TTTGATCTGTCAATATGATGTAGTGTCTTATTT-----TGCTTGTGTAG 317683
Qy 1182 TGTCTG 1241
Db 317682 TTGCAGTTTCTTCTAGCCGCTGATGGTCTTTTCCAAATTTGCAATTTTGTGAGTGGCTGGT 317623
Qy 1242 TATTATTCATATTTCTTGAATGTGGGTAAACATCTTTAGATTTGAAGTTTCTCTAGACC 1301
Db 317622 ACTAGTTGTTCTTCTCT-----GTTTAGTGTCTTCTTCTAGGAGC 317582
Qy 1302 TCTTTTAGGCTGATTTGAAGATAGATATTTCTTTTACATCTGATTTTATCTTGAAGTGTCT 1361
Db 317581 TCTTGTAGGCGAGGCTGTTGGTGACAAAATCTCTCAGCAATTTGCTGTCTGTAAAGTAT 317522
Qy 1362 TTTCTTCTCTCAACTATTGTGACAGAAAGTTTCTTAAGTCAGTACTCTGGCCTGACAT 1421
Db 317521 TTTATTTCTCTTCACTTATGAGCTTAGTTGGCTGGATATGAGATTTCTGGGTGAAA 317462
Qy 1422 CTGTAGTCTCTTGGAGTCTGTAGACATCTGTGAGGCGCTTCTTACATTTTGAAGTTTCT 1481
Db 317461 TTTCTTTCTTTAAGAAATGTAAATATCGGCCCCCACTCTCTCTCTGCTGTGTAGAGTTTCT 317402

Qy 1482 ATTGGAAGTCAAGTGTAAATCTTAATACATCTGCTTTTATATGTAAATGTGCTTTTTT 1541
Db 317401 -GCCAAGAGATCCGCTGTAGTCTGATGGCTTCCCTTTGTGGTAACCCAATC-TTCTCT 317344
Qy 1542 CCCTTGCACTTTTAAATATTTCTTTTCTTGTCTTATACATTTTAGTGAATTTATATGTC 1601
Db 317343 CTCGTGCTCCCTTAAATTTTCTTCAATTTCAACTTTGGTGAATCTGCAATATGT 317284
Qy 1602 ACTGTGGGAGTTTCTTTTCCGGTCCAAATCTATTGGTGTTTTGTATGCTTTCTGTACCT 1661
Db 317283 GTCTTGGAGTGTCTCTCTCAAGGAGTATCTTTGTGTGTCTCTCTGATTTTCTGAAATGT 317224
Qy 1662 TGATAGGCACTCTTCTCAAGGTTAGGAAATTTTCTTTTGGTCTTTCTTGAATAAT 1721
Db 317223 GAATGTGCTGCTCTGATGTTGGGAAAGTCTCTCTGGAAATATCTCTGAGAGTGT 317164
Qy 1722 TTTTCC-----CTGCTTTTGAACCTGCTCTTCTTCCCTTCTCTATCTCTTTG 1767
Db 317163 TTTCCAACTTGGTCCATTTCTCCCATCATCTTCCAGTACACCAATCAGACGTAGATTG 317104
Qy 1768 GTTTTGGATAGTGTCTCTGGCTTCTGGATGTTTATGCTGCGTGAATTTTAGACTTAA 1827
Db 317103 GTCTTTTCAATAGTCCCATATTTCTTGGAGGCTTTGTTCATTTCTTTTCTTCTT 317044
Qy 1828 CATTTTCTTTGACCAAGGTATCCATTTCTTCTATCTTGTCTTCTGCTGAGATTTCTCT 1887
Db 317043 CTCGAACTTCTTCTCTCACTTCATTTCAATTTCAATTTGATTCCTCATCTGATACTTT 316984
Qy 1888 CTCTATCTTCTGTATTTCTGTCAGTGAAGCTTGT 1921
Db 316983 CTTCAGTTGATTAATCAGTACTGAAGCTTGT 316950

RESULT 12
US-08-772-440-33
; Sequence 33, Application US/08772440
; Patent No. 6046158
; GENERAL INFORMATION:
; APPLICANT: Ariizumi, Kiyoshi
; APPLICANT: Takasima, Akira
; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
; TITLE OF INVENTION: LECTINS, DEXTIN-1 AND DEXTIN-2; COMPOSITIONS AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/772,440
; FILING DATE: CONCURRENTLY HEREWITH
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTXD:493
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10409 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

```
TOPOLOGY: linear
FEATURE:
NAME/KEY: modified_base
LOCATION: 6510
OTHER INFORMATION: /mod_base= OTHER
OTHER INFORMATION: /note= "D = A or G or T"
FEATURE:
NAME/KEY: modified_base
LOCATION: 3406..6470
OTHER INFORMATION: /mod_base= OTHER
OTHER INFORMATION: /note= "K = G or T"
FEATURE:
NAME/KEY: modified_base
LOCATION: 3564..7896
OTHER INFORMATION: /mod_base= OTHER
OTHER INFORMATION: /note= "M = A or C"
FEATURE:
NAME/KEY: modified_base
LOCATION: 3497..3607
OTHER INFORMATION: /mod_base= OTHER
OTHER INFORMATION: /note= "N = A or C or G or T"
FEATURE:
NAME/KEY: modified_base
LOCATION: 3479..6422
OTHER INFORMATION: /mod_base= OTHER
OTHER INFORMATION: /note= "R = A or G"
FEATURE:
NAME/KEY: modified_base
LOCATION: 3405..6871
OTHER INFORMATION: /mod_base= OTHER
OTHER INFORMATION: /note= "S = C or G"
FEATURE:
NAME/KEY: modified_base
LOCATION: 3457..9998
OTHER INFORMATION: /mod_base= OTHER
OTHER INFORMATION: /note= "W = A or T"
FEATURE:
NAME/KEY: modified_base
LOCATION: 3595..9999
OTHER INFORMATION: /mod_base= OTHER
OTHER INFORMATION: /note= "Y = C or T"
US-08-772-440-33
Query Match 11.1%; Score 251.899; DB 1; Length 10409;
Best Local Similarity 68.0%; Pred. No. 30;
Matches 463; Conservative 5; Mismatches 176; Indels 37; Gaps 9;
QY 1385 AGAAGTTTTTTCTAAGTGCAGTAGTCTGGCCGTGACATCTGTAGTCTCTTGGAGTCTGTAG 1444
DB 4097 AAAATAATTAATCTGGCTGTAGCAGTCTGGGCTCATATCTGTGTCTTCTAGAGTTTGTAG 4156
QY 1445 CACATCTGTGAGGCGCTTCTTACATTTGAGTTTCTATTGGAAAGTCAGGTGTAATTC 1504
DB 4157 AACATCTGTTCAGGCCCTCTGGCTTTTACAGTCTTCAAT-GAGAGTCAAGCAATTAATTC 4215
QY 1505 TAATACATCTGCCTTTATATGTTAAATGGTCTTTTCCCTTGCATCTTTTAATATATCTT 1564
DB 4216 TAGTAGGTTTATATATATGTTACTCAGTC-TTTCCTCTCTGCAGCTTTCACATCTT 4274
QY 1565 TCTTGTCTTACTTTTATGATTTGATTTATATGACATCTGGGAGTCTTCTTCCGG 1624
DB 4275 TCTTGTCTGTATGTTTATGCAATTTGATTTGTTGTTGGGAAATATCTTCTCTGA 4334
QY 1625 TCCATCTATTTGGTGTGTTTGTATGCTCTTGTACCTTGTAGGCACTCTTCTCTCAAG 1684
DB 4335 TCCAGTCCATTTGGTGTCTGCAATGTTTGTACCTTGATTAATCAATCTC---CTTAAGG 4391
QY 1685 TTAGGAATTTTCTTTTGGTGTCTTGAATAATTTTCCCTGCTTTTGACCT--GCC 1742
DB 4392 TTGAAGACATTTCTTTTATGATTTGTTGAAATAATTTCTGTGCAATTTAACTTGGCY 4451
QY 1743 TTCTTCCCTTCCCTTATCTCTTTGG-----TTTTTGCATAGTGTCTC 1785
```

```
4452 TTCTTTTCTTGTCTATTCCYATAGACTAGTTTGTGTTGTTGTTGTTTAACTGT 4511
1786 TGGCTTCTCGGATGTTTATGCC---TGGATTATTTAGACTTAACATTTTCTTTGAC- 1840
4512 AGAATTCCTGGAGTTTGTGCGCTKGAGTTTCTTTTWTATTAATATTTTCITKGACT 4571
1841 CAAGGTATCCATTTCTTCTATCTTCTTCACTGCTGAGATCTCTCTCTCTAATCTTCT 1900
4572 GGAGATAYWCCTTCTTCTATCTCTCTTCACTGCTGAGTTTCTCTTCTTCAATTCCTA 4631
1901 TATT---CTGTCACTGAGGCTTGTCTCTGAGGTTCTCTGTTGGTCTCTTAATTTTCAAT 1957
4632 TACTGAGTTGTTGGTGAGGCACTCTTCTAGACCTTGTGTTCTTCCCAATATACAT 4691
1958 TCCAGATTTCTTCACTGTTGGGTTTGTGTTTATTAATCTATTTCCACATTTCCAGTCTGA 2017
4692 TCCAG---TTTCATGCTGATTTTACTTATCGATTATATTTTACTTTCATGACATAA 4746
2018 AATGTTTACTCATTTTCTC 2038
4747 CTCCTTTTCATTTATTTATTC 4767
```

```
RESULT 13
US-09-426-290-1
; Sequence 1, Application US/09426290
; Patent No. 6410712
; GENERAL INFORMATION:
; APPLICANT: Berglind Ran Olafsdottir
; APPLICANT: Jeffrey Gulcher
; TITLE OF INVENTION: HUMAN NARCOLEPSY GENE
; FILE REFERENCE: 2345, 2001-000
; CURRENT APPLICATION NUMBER: US/09/426,290
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 168575
; TYPE: DNA
; ORGANISM: Homo Sapiens
; NAME/KEY: CDS
; LOCATION: (21181)...(21403)
; NAME/KEY: CDS
; LOCATION: (95252)...(95430)
; NAME/KEY: CDS
; LOCATION: (101753)...(101996)
; NAME/KEY: CDS
; LOCATION: (110324)...(110439)
; NAME/KEY: CDS
; LOCATION: (124058)...(124278)
; NAME/KEY: CDS
; LOCATION: (127009)...(127130)
; NAME/KEY: CDS
; LOCATION: (128910)...(129139)
; US-09-426-290-1
Query Match 10.9%; Score 248.997; DB 1; Length 168575;
Best Local Similarity 51.8%; Pred. No. 2;
Matches 857; Conservative 0; Mismatches 695; Indels 102; Gaps 14;
QY 292 TCTATTCTTGTATTTCTATCTGGCTCATTTTAACTCAGTAGTGTGTTGTTTCC 351
DB 156020 TTTTGCCTTCAATTTTGTATGACCTAGTAGTCAATTCAGGACAGGTGTTTCC 156079
QY 352 ATAAGTTTCTAAGTTTCTGTTGTTTCTGTTGTTTCTATCTAGATTAAAGCTGT 411
DB 156080 ATGTAGTTGAGGAGCTTTTGAGTGAGT---TCTTAACTCCTGAGTTCTAGTTGAC 156136
QY 412 GGTGTCAGATAGGACATAGATATTTTCAATTTGTCTTTTATCTGTCGAGACTTGT 471
DB 156137 CGTGGTCTGAGACAGATTTGTTATATATCTGATCTTATATCTGCTGAGGAGCTT 156196
```

QY 472 TGTGTTGAATATGTAATCAATTTTGA--GAGTTTCATAGGGTGTGACAGAGAGTAC 529
Db 156197 TACTTCCAACTATGTGGTCAATTTTGAATAGGTGTGGTGTGGTGTGAGAGAAATGAT 156256
QY 530 AGTCT-TTGCGTTTGGTGAATAGTCTGTGAATATCTCTAGGTCCTAGTGTGTTATGAC 588
Db 156257 ATTCTGTGATTCGGGTGGAGAGTCTGTAGATGTCTATTAGTCTGCTTGGTGCAGAG 156316
QY 589 ATCAGTTAGTCCAGCAATTCCTGTGTTCTGTTTGGTGTGATGACCTAACCTGTGGAG 648
Db 156317 CTGAGTTCAATTCCTGGATATCCCTGTGTAACCTCTCTGCTGTGATCTGCTTATATTG 156376
QY 649 AGAATGGGGTATTGAAGTAGCCCACTATCTGTGTGT-GAGGTCAATATGTGATTTAGCT 707
Db 156377 ACAGTGGGGTGTAAAGTCTCCCAATTATTATGTGTGTGTAGTCTAAGTCTCTTTGTAGT 156436
QY 708 CTAGCTGTGCTGTTTATGAACCTGGGTGACATGTGTTTGGTGATAGACATTAAGAA 767
Db 156437 CACTCAGAGCTGTGTTTATGAATCTGGTGTCTCTATATTTGGGTGCATATATATTAGGA 156496
QY 768 TTGCAATGCTCTCTGTGGTGAATTT-TCTTTGATGCTCTATGTAGTATTTCTCCCAATCTC 826
Db 156497 TAGTTAGCTCTCTGTGTTCAATGATCCCTTACCATATGTAAAGCCCTCTTTGTCTC 156556
QY 827 ATCTGCTTGTGTTTGGGTTAAGTCT-ATTAGTCAGATATTAATAAGACTGTATCGGCTT 885
Db 156557 TTTTGAICTTTGTGGTTAAAGTCTGTTTATCAGAGACTAGGATTCGAACCCCTGCCT 156616
QY 886 GCTTCTTAGGGCCATTTGCTAGATATCTTT-TTCCATCTTTTACTCTAAGGTGATGC 944
Db 156617 TTTTGTGTTTCCATTTGCTTGGTAGATCTTCCCTCCATCTTTTACTTTGAGCCTATGTG 156676
QY 945 TATCCATG--GTAGGTGTCTTTTGGATGACAGCAGTAGGATGATCTTGTGTTTCATA 1001
Db 156677 TGCTCTGCAGTGTGATGGTCTCTGAAATACAGCACACTGATGGTCTGTGACTCTTTA 156736
QY 1002 TCCATCTGTGTACCCAGTATCTTTTCTAGAGAAATTAAGATCAATGAGTCATGTGTT 1061
Db 156737 TCCAAATTTGCCAGTCTGTGCTTTTAAATGGAGCAATTAGTCCCTT----- 156782
QY 1062 GAGAATTATCAATCAGCAGTGTGTTGTGATCTTCTGTTATCTTGCACTTGTGAAGTGTG 1121
Db 156783 -----TACATTTAAAGTTAATA 156799
QY 1122 TG 1181
Db 156800 TTGTTATGTGTGAATTTGATCCTGTCATTTGA-----ATGTTAGTGTGTTATT 156848
QY 1182 TGTCTGT 1241
Db 156849 TGTGTTGT 156908
QY 1242 TATTATTCATATTTCTTGAATGTGGGTAAACATCTTTTAGATTTGAAGTTTCTCTAGCC 1301
Db 156909 TTGCAGTGGCTGTACTGTTGTTCTTCCATGTTTAG-----TGCTTCTTCAGAGC 156963
QY 1302 TCTTTTAGTCTGCAATTTGAAGATAGATATTTCTTTTACATGATGATTTTATCTAGAGTCT 1361
Db 156964 TCTTTTAGGGCAGCCCTAGTGTGTGACAAATTTCTCAGCAATTTGCTGTCTGTAAGGAT 157023
QY 1362 TTTCTTTCTCCAACTATTGTGACAGAAAGTTTCTTAAGTGCAGTGTGCTGGCTGACAT 1421
Db 157024 TTTATTTCTCTTCACTATGAGCTTAGTTTGGCTGATATGAAATCTGGGTGAA 157083
QY 1422 CTGTAGTCTCTTGGAGTCTGAGCACATCTGTGTGAGGGCCCTCTTACATTTTGAAGTTTCT 1481
Db 157084 TTTCTTTCTTTAAGAATGTTGAATTTGGGCCCCACCTCTCTCTGACTGTGAGAGTTCT 157143
QY 1482 ATTGGAAGTCAAGTGTAAATCTTAATACATCTGCCCTTATATGTTAATTTGGTCTTTT 1541
Db 157144 CGCAGAGA-TCCGCTGTGTGTGTGGGCTTCCCTTTTGGGG-TAACCCGACCTTTCT 157201
QY 1542 CCTTGCATCTTTAAATATCTTTCTTTGTTCTATACCTTTTAGTGAATTTGATTATTATGC 1601

RESULT 14

US-09-818-512-3

; Sequence 3, Application US/09818512

; Patent No. 6537780

; GENERAL INFORMATION:

; APPLICANT: BEASLEY, Ellen et al.

; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES

; FILE REFERENCE: CL001192

; CURRENT FILING DATE: 2001-03-28

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 116592

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1)...(116592)

; OTHER INFORMATION: n = A,T,C or G

US-09-818-512-3

Query Match

10.7%; Score 242.997; DB 1; Length 116592;

Best Local Similarity 51.2%; Pred. No. 3.1;

Matches 847; Conservative 0; Mismatches 705; Indels 102; Gaps 13;

QY 292 TCTATTTCTGATTTCTATCTTGGCTCATTTTAACTAGTAGTAGTGTGTTGTTTCC 351

Db 51756 TTTCTGCCCTTCATTTCTGTTATGTACCCAGTAGCTATTCCAGGAGCAGGTTGTTTCAGTTTCC 51815

QY 352 ATAAGTTTCTAAAGTTTCTGTTGTTTCTGTTGTTGTTTATCTAGATTTAAGCTGT 411

Db 51816 ATGTAGTTGAGCG---GCTTTGAGTGAGATTTCTTAATCTGAGTCTTAGTTGATTCAC 51872

QY 412 GGTGTTCAGATAGGACATAGATATTATTCAATTTGTCTTTTATCTGTGAGAGCTTGTCT 471

Db 51873 TGTGGTCTCAGAGACAGATTTGTTATAAATTTCTGTTCTTTTACATTTTGTGAGGAGCTT 51932

QY 472 TGTGTTGAATATGATTCATCAATTTTGA--GAGTTTCATAGGCTCTGACAGAGGTAC 529

Db 51933 TACTTCCAACTATGTGGTCAATTTTGAATAGGTGTGGTGTGCTGAAAAAATGTAT 51992

QY 530 AGTCT-TTGTGTTTGGTGAATAGTCTGTAAATATCTTAGGTCACCTTGTGTTTATGAC 588

Db 51993 ATTCTGTTGATTGGGGTGGAGAGTCTGTAGATGTCATTTAGTCTGCTTGTGGCAGAG 52052
QY 589 ATCAGTTAGCTCCAGCATTTCTCTGTTTCGTTTCTGTTGAGATGACCTAACTGTTGGAG 648
Db 52053 CTGAGTCAATTCCTCGGGTATCCCTTGTGACTTTCTGTCGTGTGATCTGCTCAATATG 52112
QY 649 AGAATGGGGTATTGAAGTAGCCCACTATCTGTGTGTGAG-GTCAATATGTGATTTTAGCT 707
Db 52113 ACAGTGGGGTGTAAAGCTTCCATTAATATGTTGGAGTCTAAGTCTCTTTTAGGT 52172
QY 708 GTAGCTGCTGTTTATGAACTTGGGTGACATGTGTTTGGGATAGACATTAAGAA 767
Db 52173 CACTGAGGACTTGTCTTATGAATCTGGGTGCTCTGTATTTGGGTGCAATAATATTAGGA 52232
QY 768 TTGCAATGCTCTCTTGTGGGATTT-TCTTTGATGCTATGTAGTATCTTTCCCAATCTC 826
Db 52233 TAGTTAGCTCTCTTGTGAATGATCCCTTACATTAATGTAATGGCTTCTTTGTCTC 52292
QY 827 ATCTGCTTAGTTTGGGTTTAAAGTCT-ATTAGTCAGATATTAAATGACGTGATPCGGCTT 885
Db 52293 TTTTGATCTTTGTTGGTTTAAAGTCTGTTTATCAGAGACTAGGATTTGCAACCCCTGCT 52352
QY 886 GTTCTTAGGGCCATTTGCTTAGATATCTT-TTCCATCTCTTTACTCTAAGGTGATGC 944
Db 52353 TTTTGTGTTTCCATTTGCTTGGTAGATCTTCTCCATCTCTTTTATTTTGGAGCTATGTG 52412
QY 945 TATGCATG--GTAGTTGTCCTTTTGGATGACAGCTAGTAGTATCTTGTGTTTTCATA 1001
Db 52413 TGTCTGACAGTGGAGTGGTTCCTGAAATACAGACACTGATGGGTCTGACTCTTTA 52472
QY 1002 TCCATCTGTACCCAGATCTTTTCTAGAGAAATTAAGATCAATGAGTATGATGTT 1061
Db 52473 TCCACCTTGCCAGCTGTGTCTTTTAAATGAGAAATTTAG-----TCCATTTATATT 52524
QY 1062 GAGAAATATCAATGAGCAGTGTGTTGGATCTGTTTATCTCTGACATGTTGAAAGTGTG 1121
Db 52525 TAAAGTTAATATTGTTATGTTGAAATTTGATCTCTGATTAATGATGTTAGCTGTGATTT 52584
QY 1122 TGT 1181
Db 52585 TGCTCATTTAGTTGATGAGTTTCTTCTAGTCTGATGCTTTTACATTTTGGCATGATT 52644
QY 1182 TGTCTGT 1241
Db 52645 TTGAGCGGCTGGTACCGGTTGTTCTTCTC----- 52675
QY 1242 TATTATTCATATTTTCTTGAATGTGGGTAAACATCTTTAGATTGAAGTTTTCCTTAGCC 1301
Db 52676 -----ATGTTTAGCGCTTCTTCCAGGAGC 52699
QY 1302 TCTTTTAGCTCGCATTTGAAGATAGATATCTTTTACATCTGATTTTATCTTAGAATGTC 1361
Db 52700 TCTTTTAGGGAGGCGCTGGTGGTACAAAATCTCTCAGCATTTGCTGTCTATAAGTAT 52759
QY 1362 TTCTTTCTCACAATTTGTGACAGAAAGTTTCTTAAGTGCAGTAGCTGCGCTCGACAT 1421
Db 52760 TTTATTTCTCTTCACTATGAGCTTAGTTTGGCTGGATAGAAATCTGGGTGAAAT 52819
QY 1422 CTGTAGTCTCTTGGAGCTGTAGACATCTGTGTGAGGGCTTCTTACATTTGAGTTTCT 1481
Db 52820 TTTCTTTCTTTAAGAAATGTTGAATATTGGCCCACTCTCTCTCTGCTTGTAGGGTTTCT 52879
QY 1482 ATTGAAAGTCAGGTGTAATTTCTAAATACATCTGCTTTATCTTAAATGTTCTTTTT 1541
Db 52880 GCCGAGAG-TCCGCTGTAGTCTGATGGGCTTTTCTTTG-AGGTTAACCCGACCTTTCT 52937
QY 1542 CCGTTGCACTTTTAAATCTTTCTTTGTTTCTATCTTTTGTAGTTTGAATTAATGTC 1601
Db 52938 CTCCTGGCTGCCCTTAAACATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 52997
QY 1602 ACTGTGGGAGTTCTTTTCTGGTCCCAATCTATTTGGTGTGTTTGTATGCTCTTGTACCT 1661

Db 52998 GTCTGGAGTTGCTCTCTCTCGAGGAGTATCTTTGTGGGCTTCTCTGTATTTCCCTGAATCT 53057
QY 1662 TGATAGGCATCTCTTTCTCAAGTTTAGGAAATTTTCTTTTGTGTTTCTTTGAAATAT 1721
Db 53058 GAACGTTGGCCTGCTTGTAGATTGGGAGTCTCTCTGGATATATCTCTGCAGAGTGT 53117
QY 1722 TTTC-----CTGCTTTTACCTGCTTCTTCCCTTCTCTTCTTCTTCTTCTTCTTCT 1767
Db 53118 TTTCCAACTTGGTTCCATTTCTCCACATCACTTTTCAAGTATACCAATCAGAGCTAGATTG 53177
QY 1768 GTTTTTCATAGTCTCTGCTTCTCTGATGTTTTATGCTTGGATTTATTTAGACTTAA 1827
Db 53178 GTCTTTTTCATAGTCCCATATTTCTTGGAGCTTTGCTCACTTTCTTTTATCTTTTCT 53237
QY 1828 CATTTTCTTTTGAACCAAGTATCCATTTCTTCTATCTTCTTCTTCTTCTTCTTCTTCT 1887
Db 53238 CTCTAACTTCTCTTCTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 53297
QY 1888 CTCTATCTCTTGTATTTCTGCTGAGTGAAGCTTGT 1921
Db 53298 CTTCAGTTGATCGCATCGGCTCTCTGAGGCTTCT 53331
RESULT 15
US-09-751-389-3
; Sequence 3, Application US/09751389
; Patent No. 6630334
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO01067
; CURRENT APPLICATION NUMBER: US/09/751,389
; CURRENT FILING DATE: 2001-01-02
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 786431
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(786431)
; OTHER INFORMATION: n = A,T,C or G
US-09-751-389-3

Query Match 10.5%; Score 239.897; DB 1; Length 786431;
Best Local Similarity 51.4%; Pred. No. 0.51;
Matches 833; Conservative 0; Mismatches 686; Indels 103; Gaps 13;
QY 324 TAATCAGTAGTGTGTTGGTTTCCATAAGTTTGTAAAGTTTCTGTTTCTGTTCTGTTG 383
Db 580096 TCATTACAGGAGGCTGTTTCTGTTTCTTCCATGATGTTTGGTTTTCAGTGTGAGT 580151
QY 384 TTGTTGTTGTTTCTAGATTTTAAAGCTGTGGTGTGAGTGTGAGTGTGAGTGTGAGTGT 443
Db 580152 TTAATCTCTGAGCTCTAGTTTGTATGTTCTGAGTGTGAGTGTGAGTGTGAGTGT 580211
QY 444 ATGTCTTTTATCTGTCAGACTTCTTGTGTTTGAATATGTTCAATTTTGA--GA 501
Db 580212 ATTCTTTTACATTTCTGCTGAGGAGTCTTTACTTCAACTGTGTGCTCAATTTTGAATAG 580271
QY 502 GTTTCATAGGGTGTGTCACAAAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 560
Db 580272 GTGCTGTGTGGTGTGAAAGAAATGATATCTGTTGATTTGGGTGTGAGAGTCTGTAG 580331
QY 561 ATATCTTAGGTCCACTTGTGTTTATGACATGATGATGATGATGATGATGATGATGATGAT 620
Db 580332 ATGTCTATTAGGTCCACTTGTGTGAGAGTGTGATGATGATGATGATGATGATGATGAT 580391
QY 621 TTTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 680

Db 15727 ACAGTGGGGTAAAGTCTCCCATTAATTAATGTGTGGAGCTAAAGTCTCTTTGTAGGT 15786
QY 708 GTAGCTGTCTGTTTATGAACCTTGGGTGACATGTGTTGGTGCATAGACATTAAGAA 767
Db 15787 CACTGAGGACTTGCCTTATGAATCTGGGTGCTCTGTATTTGGGTGCATAAATATTTAGGA 15846
QY 768 TTGCAATGCTCTTGGTGGATT--TCTTTGATGCCCTATGTAGTATTTCTTCCAAATCTC 826
Db 15847 TAGTTAGCTCTCTTGTGGAATGTATCTTACCATTAATGAATGCCCTCTTTGTGCTC 15906
QY 827 ATCTGCTAGTTTGGGTTAAGTCT-ATTAGTCAGATTAATAAGATGATCTGCTT 885
Db 15907 TTTTGTATCTTTGGTGTAAAGTCTGTATATCAGACACTAGGATGCAACCCCTGCT 15966
QY 886 GCTTCTTAGGGCCATTTGCTTAGAATATCTT-TTCCATCTCTTTTACTCTAAGGTGATGTC 944
Db 15967 TTTTGTGTTTCCATTTGGCTGTAGATCTTCCCTCATCTCTTTTATTTTGGCCATGTC 16026
QY 945 TATCCATG---GTAGTGTCTTTTGGATGCCAGTAGAGTCTGTTTTCATA 1001
Db 16027 TGTCTCTGACGTGAGTGGGTCTCTGAATACAGACACTGATGGGTCTGACTCTTTA 16086
QY 1002 TCCATTTCTGTTACCCAGTATCTTTTCTAGAGAAATTAAGATCATTCAGTCAATGATGTT 1061
Db 16087 TCCAACTTGCCAGTCTGTGCTCTTTTAAATTCAGAAATTTAG-----TCCATTTATTT 16138
QY 1062 GAGAAATTAATGAGCAGTGTGTTGGATCTTGTGTTATCTTGACATGTTGAAGTGTG 1121
Db 16139 TAAAGTAAATATGTATGTGGAATTTGATCCCTGTCAATTAAGTGTAGCTGGCGATTT 16198
QY 1122 TGT 1181
Db 16199 TCTCATTTAGTGTGAGTCTTCTTAGTCTGATGGTCTTTTACATTTTGGCATGATTT 16258
QY 1182 TGT 1241
Db 16259 TTGACGGCTGTGACCGGTGTGCTCTTCC-----16289
QY 1242 TATTATTCATATTTTCTTGAATGTGGTAAACATCTTTAGATTAAGTCTTTCTCTAGCC 1301
Db 16290 -----ATGTTTACCGCTTCTTCCAGAGC 16313
QY 1302 TTTCTTAGTCTGCAATTTGAAGATAGATATCTTTACATCTGATTTTATCTTAGAATGTC 1361
Db 16314 TCTTTTAGGGCAGGCTGCTGTGTGACAAATCTCTCAGCAFTTCTCTGTCTATAAAGTAT 16373
QY 1362 TTTCTTTCTCCAACTATTGTGACAGAAAGTCTTCTAAGTGCAGTCTCTGGCTGACAT 1421
Db 16374 TTTATTTCTCTCACTTATGAGCTTATTTTGGCTGGAATGAATCTGGGTGAAA 16433
QY 1422 CTGTAGTCTCTTGGAGTCTGTAGCAATCTGTGAGGGCTCTTTACATTTTGTGTTCT 1481
Db 16434 TTTCTTTCTTTAAGAAATCTGAAATATGGGCCCACTCTCTCTGCGCTTGTAGGGTTCT 16493
QY 1482 ATTGAAAAGTCAGGTGAATCTAAATACATCTGCTTTATATGTTAATTTGGTCTTTT 1541
Db 16494 GCCGAGAGA-TCCGCTGTAGTCTGATGGCTTTCCITTTG-AGGTAACCCGAACTTCT 16551
QY 1542 CCGTTGCAATTTTAAATCTTTCTTTGTTCTATCTATCTTTTAGTATGATTTATATGC 1601
Db 16552 CTCGGCTGCGCTTAAACATTTTCTTTCACTTCAACTTTTGGTGAATCTGACAAATATGT 16611
QY 1602 ACTGTGGGGAGTTCTTTTCCGGTCCAAATCTATTTTGGTGTGTTTGTATGCTTCTGTACCT 1661
Db 16612 GTCCTTGGAGTGTCTCTCTCGAGGAGTATCTTGTGGCGTCTCTGTATTTCTGATCT 16671
QY 1662 TGAATAGGATCTCTTTCTCAAGGTTAGGAATTTTCTTTTGGTGTGTTTCTTGAATAAT 1721
Db 16672 GAACGTTGGCTGCTCTCTAGATTTGGGAAAGTTCTCTCGGATAAATATCTGAGAGTGT 16731
QY 1722 TTTTCC-----CTGCTTTTGACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1767
Db 16732 TTTTCCAACTGTTGTTCCATCTCCCATCTACATCTTCCAGGTACACCAATCAGACGTAGATTG 16791

QY 1768 GTTTTTCATAGTGTCTCTGGCTTCTCTGGATGTTTATGCTGCTGATTTATTTAGACTTAA 1827
Db 16792 GTCITTTTCATAGTCCCATATTTCTTGGAGCTTCTGCTCATTTCTTTTATCTTTT 16851
QY 1828 CATTTTCTTTGACCAAGTATCCATTTCTTCTATCTTGTCTTCACTGCTGAGATCTCT 1887
Db 16852 CTCATAACTTCCCTTCTGCTTCAATTTCAATTTCAATTTCAATTTCAATTTCAATTT 16911
QY 1888 CTTCTATCTCTTGTATTTCTGTCAGTGGGCTTGT 1921
Db 16912 CTTCCAGTTGATCCATCGGCTCTCTGAGCTTCT 16945

RESULT 17
US-09-340-620A-63
; Sequence 63, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; CURRENT FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 32042
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-340-620A-63

Query Match 10.58; Score 239.797; DB 1; Length 32042;
Best Local Similarity 51.18; Pred. No. 11;
Matches 845; Conservative 0; Mismatches 707; Indels 102; Gaps 13;
QY 292 TCTATTTCTGATTTCTATCTTGGCTCATTTTAACTCAGTGTGTTGTTTCC 351
Db 15370 TTTCTGCTTCATTTCTGTTATGTACCCAGTAGTCAITTCAGGACAGGTTGTTTCC 15429
QY 352 ATAAGTTGTAAGTTTCTGTTTCTGTTTGTGTTGTTTGTATCTAGATTTAAGCTGT 411
Db 15430 ATGTAGTTGAGCG--GCTTGTAGTGTAGTCTTAATCTGAGTTCTAGTTGATGAC 15486
QY 412 GGTGCTCAGATAGACATAGAGTATTTTCAATGTCTTTTATCTGTGTGAGACTTGTCT 471
Db 15487 TGTGCTCTGAGAGATAGTTTGTATATAATTTCTGTTTACATTTTCTGCTGAGAGCTT 15546
QY 472 TGTTTTGAATATGATTTCAATTTTGA--GAGTTTCATAGGCTGCTGACAAAGGTAC 529
Db 15547 TACTTCCACTATGTGTGTCATTTTGAATAGTGTGTTGCTGTGCTGAAATAATGAT 15606
QY 530 AGTCT-TTGTGTTTGTGTAATAGTCTGTAATAATCTCTAGTCCACTTGGTTTATGAC 588
Db 15607 ATTCTGTGTTTGGGTGGAGAGTTCTGTAGATGTCATTAGGTCTGCTGTGTCAGAG 15666
QY 589 ATCAGTTAGCTCCAGCATTTCTCTGTTTCTGTTTGTGTTGTTGAGATGACCTAACTGTTGAG 648
Db 15667 CTGAGTTCAATCTCTGGGATCTTGTGTGATTTCTGTCTCATTTGATCTGTCTAATGTTG 15726
QY 649 AGAATGGGGATTGAAGTACCCACTATCTGTGTGTGAG-GTCAATATGTGATTTTGTAGCT 707
Db 15727 ACAGTGGGGTAAAGTCTCCCATTTAATGTGTGGGAGTCTAAGTCTCTTTTGTAGCT 15786
QY 708 GTAGCTGTGCTGTTTGTATTAACACTTGGGTGACATTTGTGTTGGTGCATAGACATTAAGAA 767

Db 15787 CACTGAGACATGCTTTAATGAACTGGGTGCTCTGTATGCGTCAATAAATATTAGGA 15946
Qy 768 TTGCAATGTCCTCTGGTGGATTT-TCCCTTTGATGCTATGATGATTTCTTCCCAATCTC 826
Db 15847 TAGTTAGCTCCTCTGTTGAAATGATCCCTTTACCAATATGATGAGGCTTCTTTGCTC 15906
Qy 827 ATCTGCTTAGTTTGGGTTTAAGTCT-ATTAGTCAGATATTTAAATGACGTATTCGGCTT 885
Db 15907 TTTTGAATCTTTGTTGTTTAAAGTCTGTTTATCAGAGACTAGGATTCGAACCCCTGCT 15966
Qy 886 GCTCTTAGGGCCATTTCTTAGAATATCTT-TTCCATCCTTTTACCTAGGTGATGTC 944
Db 15967 TTTTGTGTTTCCATGCTGATGATCTTCCCTCCATCCTTTATTTAGAGCTATGTC 16026
Qy 945 TATCCATG---GTAGTTGCTTTTGGATGACAGCTAGGATGATCTTGTGTTTCATA 1001
Db 16027 TGCTCTGACGCTGAGATGGGTTTCCGATACAGCACACTGATGGTCTTGATCTCTTA 16086
Qy 1002 TCCATCTGTTACCCAGTATCTTTTCTAGAGAAATTAAGATCATGAGTCATGATGTT 1061
Db 16087 TCCAACTTGCAGTCTGTGCTTTTAAATGAGAAATTTAG-----TCCATTTATATT 16138
Qy 1062 GAGAAATATCAATGAGCAGTGTGTTGTAATCTTGTGTTTATCTTGCACTTGTGAAGTGTG 1121
Db 16139 TAAAGTTAAATTTGTTAGTGAATTTGATCCTGTCTATGATGATGTTAGCTGCGATTT 16198
Qy 1122 TGT 1181
Db 16199 TGCTCATAGTGTGAGTGTCTTCTCTAGTCTGATGCTTATACATTTTGGCATGATT 16258
Qy 1182 TGTCTGTGCTGT 1241
Db 16259 TTGAGGGCTGTGTACCGGTTGTTCCCTTCC-----ATGTTTACCCTTCTTCAGGAGC 16313
Qy 1242 TATTATTCATATTTCTTGAATGGGTAAACATCTTTAGATTTGAAGTTTCTCTAGACC 1301
Db 16290 -----ATGTTTACCCTTCTTCAGGAGC 16313
Qy 1302 TTCTTTAGGCTGCAATTTGAAGATAGATATCTTTACATCTGATTTATCTTGAATGTC 1361
Db 16314 TCTTTTAGGCGAGCCCTGGTGGTACAAATCTCTAGCAATTTGCTGTCTATAAAGTAT 16373
Qy 1362 TTCTTTCTCCAACTATTTGACAGAAAGTTTCTTAAGTGCAGTGTCTGGCTGCACAT 1421
Db 16374 TTTATTTCTCCTTACATTTAGAGCTTAGTTGGCTGGATGAATTTCTGGTTCGAAA 16433
Qy 1422 CTGTAGTCTCTGGAGTCTGACACATCTGTGAGGGCTTCTTACATTTTGAATTTCT 1481
Db 16434 TTCTTTCTTTAAGATGTTGAATATGSCCCCACTCTCTCTGCTGTAGGGTTCT 16493
Qy 1482 ATTGGAAGTCAAGTGTAAATCTTATACATCTGCTTTATATGTTAATGCTTTTCT 1541
Db 16494 GCGAGAGA-TCCGCTGTAGTCTGATGGCTTCTCTTTG-AGGTTAACCCGAACTTCT 16551
Qy 1542 CCCTTGATCTTTTAAATTTCTTTGTTGTTCTATATCTTTAGTGAATTTGATTTATGC 1601
Db 16552 CTCTGGCTGCCCTTAACTTTTCTCTTCAATTTGGTGAATCTGACATTTATCT 16611
Qy 1602 ACTGTGGAGTCTTTTCCGGTCCAACTATTTGGTGTGTTTGTATGCTTCTGTAACCT 1661
Db 16612 GTCTTGAGTGTCTCTCTGAGAGTATCTTTGTGGCGTCTCTGTATTTCTCTGAATCT 16671
Qy 1662 TGATAGGCATCTCTTTCTCAAGTGTAGGAAATTTTCTTTTGGTGTGTTTCTTGAATAAT 1721
Db 16672 GAACGTGGCTGCCCTTGTAGATTTGGGAGTGTCTCTCGATATATCTCTGAGATGT 16731
Qy 1722 TTTTCC-----CTGCTTTGACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1767
Db 16732 TTTTCCAACTGGTTCATTTCTCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 16791
Qy 1768 GTTTTTCATAGTGTCTGCTGCTCTGATGTTTATGCTGCTGATTTTATAGACTTAA 1827

Db 16792 GTCTTTTACATAGTCCCATATTTCTTGGAGGCTTTGCTCATTTCTTTTCTTTT 16851
Qy 1828 CATTTCTTTGACCAAGGTATCCATTTCTTCTATCTTGTCTTCACTGCTGAGATCTCT 1887
Db 16852 CTCTAAACTTCCCTTCTCGCTTCATTTCAATTCATTTCACTTCTGCTGATACCTTT 16911
Qy 1888 CTCTATCTCTTCTTATTTCTGCTAGTCAGGCTGT 1921
Db 16912 CTTCAGTTGATCGCATCGGCTCCTGAGGCTTCT 16945

RESULT 18
US-08-742-185-102
; Sequence 102, Application US/08742185
; Patent No. 6020476
; GENERAL INFORMATION:
; APPLICANT: Page, David C.
; APPLICANT: Reijo, Renee
; APPLICANT: Saxena, Richa
; APPLICANT: Hawkins, Trevor
; APPLICANT: Reeve, Mary Pat
; TITLE OF INVENTION: DAZ: A GENE FAMILY ASSOCIATED WITH AZOOSPERMIA
; NUMBER OF SEQUENCES: 102
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,185
; FILING DATE: 30-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/690,734
; FILING DATE: 31-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/310,429
; FILING DATE: 22-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: WHI94-07A2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 102:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40328 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-742-185-102

Query Match
Best Local Similarity 10.0%; Score 227.997; DB 1; Length 40328;
Matches 882; Conservative 0; Mismatches 785; Indels 106; Gaps 13;

Qy 292 TCTATTTCTTGAATTTCTATCTTGGCTCATTTTAACTCAGTAGTGTGTTGTTGTTCC 351
Db 19589 TTTCTGCTTTCATTTGTTATGTACCCAGTAGTCATTCAGGAGCAGGTGTTCAGTTCC 19648
Qy 352 ATAGTTTGAAGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 411
Db 19649 ATGAGTTGAGTGGTTT---TGAGTGAGATTTCTTAATCCTGAGTTCTAGTTGTCAC 19705

412 GGTGTCAGATAGGACATAGATATTAATCAATGTCCTTTATCTGTGCGAGACTTGCCTT 471
19706 TGTGGTCTCAGAGATAGT-TGGTAAATTCGTCTTTTACATCCGTCAGGAGAGCTT 19765
472 TGTGTTGAATATGATTAATCAATTTTGA--GAGTTTCATAGGTCCTGCAAGAGAGTAC 529
19766 TACTTCCCAGTATG-GGTGAGTTTTTGAATAGGTGTGTGGTGTGCGAATAAATGTA 19825
530 AGTCCTTGTGTTTT--GGTGAATAGTCTGTAATATCTCTAGTCCACTTGGTTATGA 587
19826 TATTCCTGTTGATTTGGGGTGGAGAGTCTGTAGATGCTATTAGGTCACACTTGGTGCAGA 19885
588 CATCAGTTAGTCCAGCATTTCTGTGTTGCTTTTGTGAGATGACCTAACTGTGGA 647
19886 GCTGAGTTCAATCTCTGGGTATCCTGTGTGATCTTCTGCTGCTGATCTGTCTAATGTT 19945
648 GAGAAATGGGTTATGAAGTAGCCACTATCTGTGTGAG--GTCAATATGTGATTTTACG 706
19946 GACAGTGGGGTGTAAAGTCTCCCATTAATGATGTGGAGTCTAAGTCTCTTTGTAGG 20005
707 TGTAGTGTGCTGTGTTTATGAACHTGGTGACATGTGTTGGTGATAGACATTAAGA 766
20006 TCACCTCAGACCTGTGTTTATGAATCTTGGTGTCTCTGTGTTGGTGATATATATTAGG 20065
767 ATTGCAATGTCCTCTGTGGTGAATTT-TGCTTTGATGCTCTATGATGATCTTCTCCAACT 825
20066 ATAGTTAGCTCTCTGTTGTAATGATCCCTTACCATTATGTAATGGCCTCTTGTGCT 20125
826 CATCTGCTTAGTTTGGGTTAAGTCT-ATTAGTCAGATATTAATAGACTGTATCGGCT 884
20126 CTTTCGACTTGTGTTTAAAGTCTGTTTATCAGAGACTAGGATGGAACCCCTGCC 20185
885 TGTCTCTTAGGGCCCAATGCTTAGAATATCTT-ITCCATCCTTTTACTCTAAGGTGATGT 943
20186 TTTTGTGTTTCCATTTGCTGGTGTAGATCTTCTGTGATCTTATTTTGGAGCTATGT 20245
944 CTATCCATG---GTAGTTGCTTTTGGATGACGACGATGAGTCTGTTTTCAT 1000
20246 GTGTCCTGACGTCAGATGGTTCCTCAATACAGACACACTGATGGGTCTTGACTCTT 20305
1001 ATCCATCTGTATCCAGATATCTTTTCTAGAGAAATTAAGATCAATGACATGATGT 1060
20306 ATCCAGTTGCGAGCTGTGCTTTTAAATGGAGCAITTAG-----TCCATTTGACAT 20357
1061 TGAGAAATATCAATGACGAGTGTGTTGAGATCTTGTATCTTCCACTTGTGAAGTGT 1120
20358 TTAAGATTAATTTATGATGAATTTGATCTGTCATTTATGATGTTAGCTGGTTATT 20417
1121 GT 1180
20418 TTGCTGT 20477
1181 GT 1240
20478 TTTGACGGGCTGT 20509
1241 TTAATATTATATTTTCTTGAATGTGGGTAAATCTTTAGATGGAAGTTTCTCTGAGC 1300
20510 -----ATGTTTAGTCTCTCTCAGGAG 20532
1301 CTTCTTTAGTCTGCAATTTGAAGATAGATATCTTTTACATCTGATTTTATCTTAGAATGT 1360
20533 CTCCTATAAGGAGGCTGTGGTGTGACAAATCTCTCAGCATTTGCTGTCTGTAAAGTA 20592
1361 CTTTCTTTCTCCAACTATTGTGACAGAAAGTTTCTTAAGTGCAGTGTGTGGCTGACA 1420
20593 TTTTATTTCTCTCTGCTATGAGCTTAGTTGTGGCTGGATATGAATTTCTGGGTGAAA 20652
1421 TCTGTAGTCTCTGTAGTCTGTAGACATCTGTGACGGGCTCTTACATTTTGAAGTTTC 1480
20653 ATATTTTCTTTTAAAGATTTTGAATATTGTGGCCCCCACTCTCTCTGCTGTGAGGGTTC 20712
1481 TATTGAAAGTCAGGTGTAAATCTTAAATACATCTGCTTTTATATGTTAAATGGTCTTTT 1540

Db 20713 TGCCGAGAGATCCACTGTAGTCTGA--TGGGCTTCCCTTTGAGGGTAACCCGACCTTTC 20770
QY 1541 TCCCTTGATCTTTTAAATATTTCTTCTTCTTCTATATCTTTTCTTCTTCTTCTTCTTCT 1600
Db 20771 TCTCTGGCTGCCCTTACATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 20830
QY 1601 CACTGTGGGAGTTTCTTCTTCCGGTCCAACTCTATTGGTCTTTTGTATGCTTCTTCTTCTTCT 1660
Db 20831 TGTCTTGGAGTTGTCTTCTTCTGAGGATATCTTTGTGTGTCTTCTTCTTCTTCTTCTTCT 20890
QY 1661 TTGATAGCATCTCTTCTCAAGGTTAGGAAATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1720
Db 20891 TGAAGTTGGCTGCTTCTGATGATGAGGAGTTCTCTGATATATATCTTCTGAGAGTG 20950
QY 1721 TTTTCC-----CTGCTTTTGACCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1766
Db 20951 TTTTCCAGCTTGGTTCCATTTCTCCGATCACTTTTCAAGTACACCAAGCAGACGATGATTT 21010
QY 1767 GGTCTTTGATAGTCTCTGCTTCTGCTGATGTTTCTGCTGATTTTCTGCTGATTTTCTGATTTA 1826
Db 21011 GGTCTTTTCAATAGTCCCATATTTCTTGGAGGCTTGTCTTCTTCTTCTTCTTCTTCTTCT 21070
QY 1827 ACATTTCTTTGACCAAGGATCCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1886
Db 21071 TCTGCAAACTTCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 21130
QY 1887 TCTTCTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1943
Db 21131 TCTTCCAGTTGATGCAATGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 21190
QY 1944 TTAATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2003
Db 21191 TTGTTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 21250
QY 2004 CTTTCCAGTCTGCAAAATGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2036
Db 21251 CATCT 21283

RESULT 19
US-09-734-674-3
; Sequence 3, Application US/09734674
; Patent No. 6498022
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001018
; CURRENT APPLICATION NUMBER: US/09/734,674
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 202001
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(202001)
; OTHER INFORMATION: n = A, T, C or G
US-09-734-674-3

Query Match 9.1%; Score 208.398; DB 1; Length 202001;

Best Local Similarity 50.1%; Pred. No. 2.7;

Matches 835; Conservative 0; Mismatches 711; Indels 120; Gaps 14;

QY 252 TCTATTTCTTGAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 351

Db 145374 TTTCTGCTTCT 145433

QY 352 ATAAAGTTTGAAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 411

Db 145434 ATATAGTTGAGCGGTTTGTAGTGAGT---TCTTAATCCTGAGTCTAGTTGATGCA 145490
QY 412 GGTGGTCAGATGAGACATAGAGTATTATTTCAATGTCCTTTTATCTGTCGAGACTTCTCT 471
Db 145491 TGTGGTCGAGAGACAGTTTGTATTAATTTCTGATCTTTTACATTTGCTGAGAGAGCTT 145550
QY 472 TGTGTTGAAATATGATTAATTTTGA--GAGTTTTCATAGGGTCTGCAAGAAGGTAC 529
Db 145551 TACTTCCAACTATGTGGTCAATTTTGAATAGGTGTGGTGTGGTCTGCTGAAATAATGAT 145610
QY 530 AGTCT-TTGTGTTTGTGGAATAGTCTGTAATATCTCTAGGTCCTGCTGCTGTTATGAC 588
Db 145611 ATTCTGTTGATTTGGGGTGGAGAGTCTGTAGATGCTTATPAGGTCGCTGTGGTCAGAG 145670
QY 589 ATCAGTGTAGTCCAGCAATTTCTGTTTGTGTTTTTGTGAGATGACCTAAGTGTGGAG 648
Db 145671 CTGAGTTCAATTCCTGGGTATCCCTGTCAACTTTCTGCTGCTGATCTGCTAATGTTG 145730
QY 649 AGAATGGGTATTGAAGTAGCCCACTATCTGTGTGTGAG-GTCAATATGATTTTAGCT 707
Db 145731 ACAGTGGGATGTTAAAGTCTCCCAATTAATTTTGTGGGAGTCTAAGTCTCTTTGTAGT 145790
QY 708 GTAGCTGTGCTGTTTATGAACCTGGGTGACATTTGTTTGGTGCATAGACATTAAGAA 767
Db 145791 CACTCAGGACTTGTCTTATGAATCTGGGTGCTCTGCTATTAGATACATATATTAGGA 145850
QY 768 TTGCAATGTCTCTTTGTG-GATTTTCTGTGAGTCTATGATGATTTTCCCAATCTC 826
Db 145851 TAGTAGCTCTCTCTGTGAGTGTGATCCCTTACCAATTATGTAATGGCTT---GTCTC 145906
QY 827 ATCTGCTTAGTTTGGGTTTAAAGTCTATTAGTCAGATTAATAATGACGTATCGGCTTG 886
Db 145907 TTTTGTCTTTGTTGGTTTAAAGTCTGTT--TTATCAGAGACTATGATTCGAACCCCTGC 145964
QY 887 CTTCTTAGGCCCAATTTGCTTAGAATATCTTT-----TCCATCCCTTTTACTC 932
Db 145965 CTTTCTTTTGGTTTTTTTTTTTTTTTTTTTTTTTGGTAGATCTTCCCTCCATCCCTTTATTT 146024
QY 933 TAGGTGATCTCTATCCATG---GTAGTTGTCTTTTGGATGACAGCAGTAGATGGAT 989
Db 146025 TGAGCCTATGTGTCTCTGACGTGTGATGGTTTTTCTGAAATACAGCACACTGATGGGT 146084
QY 990 CTTGTTTTTCATCCATCTCTGTACCAGTATCTTTTCTAGAGAATTAAGATCAATGA 1049
Db 146085 CTTGACTCTTTATCCAAATTTGCCAATCTGTGCTTTT-----AATTAGACCATTCAG 146136
QY 1050 GTCAATTGATTTGAGAAATATCAATGACAGAGTGTGTTGAGATCTTTGTTATCTTGCATT 1109
Db 146137 CCATTTTACCTTTAAGGTTAATATTGTTATGTGGAATTTGATCCTGTCTATTATGATGTT 146196
QY 1110 GTGAAGT 1169
Db 146197 AGCTGGTATTGCTTTGTTACTTGTGATGAGTACTTCTTAGCATCGATGGTCTTTTACAA 146256
QY 1170 TGT 1229
Db 146257 TTTGGCATGTTTTTGCAGTGGCTGTACCAAGTTGTTCCCTCTC----- 146299
QY 1230 CTTGGAATTTATTTATTTATTTCTTGAATGTGGGTAAACATCTTAGAATTGAAGTT 1289
Db 146300 -----ATGTTTAGTGTCT 146311
QY 1290 TTTCTCCTAGCCCTTTTAGGTCGATTTGAAGATAGATATTCTTTTACATCTGATTTTA 1349
Db 146312 TCCTTCAGGAGCTCTTTTAGGCGAGCCGTGGTGTGACAAATCTCTCAGCATTTGCTTG 146371
QY 1350 TCTTAGAATGCTTTCTTTCTCCAACTAATTGTGACAGAAAGTTTTTCTAAGTGCAGTAGT 1409
Db 146372 TCTGTAAGTAATTTATTTCTCTCTCTACATATGAAGCTTAGTTGGCTGGATGAAAT 146431
QY 1410 CTGGCCTGACATCTGTAGTCTCTGTAGTCTGTAGCACATCTGTGAGGCGCTTCTTACA 1469

Db 146432 CTGGGTTGAAATTTCTTTTCTTTAAGAAATGTTGAATATTGGCCCCCACCCTCTCTCTGGCT 146491
QY 1470 TTTTGTAGTTTCTATTGGAAAGTCAGGTGTAATCTTAATACATCTGCCCTTTATATGTTAA 1529
Db 146492 TGTAGAGTTTCTGCT--GAGAGATCCGCTGTTAGTCTGATGGGCTTCCCTTTGT--GGGTAA 146549
QY 1530 TTTGGTCTTTTTCCTCTGTCATCTTTTAATATTCTTTTCTTTGTTCTATATCTTTTAGTATT 1589
Db 146550 CCCAACCTTTCTCTGCTGGCTGCCCTTAACATTTTTCCTTCAATTCACACTTTGGTGAATC 146609
QY 1590 TGATTTATGACATGTGGGGAGTTCTTTTCCGGTCCAACTATTATTTGGTGTGTTTGTATG 1649
Db 146610 TGAATAATTAATGTCTGTGGAGTTGTTATCTCGAGAGATATCTTTGTGTTCTCTCTGA 146669
QY 1650 CTTCTGTGTACCTTGATAGGCACTCTCTTCTCAAGGTTAGGAAATTTTCTTTTGGTTT 1709
Db 146670 TTTTCTGAATCTGAATGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 146729
QY 1710 TCTGAAATATTTTCC-----CTGCTTTTGACCTGCTTCTTCCCTTCC 1755
Db 146730 CTTGGGAGTGTTTTCCAACTTTGGTTCCATTTCTCCGGTCACTTTTCAGTACACCAATCG 146789
QY 1756 TCTATCTCTTCTGTTTTCATAGTGTCTCTGGCTTCTGGATGTTTATGCTCGATTA 1815
Db 146790 GACGTAGATTGTTCTTTTTCATATGCTCCATATTTCTTGGAGGCTTTGTTGTTCTT 146849
QY 1816 TTTTAGACTTAACATTTCTTTTACCAAGGTATCCATTTCTTCTATCTTCTTCTTCACTGC 1875
Db 146850 TTAATCTTTTCTCTAAACTTTCTTCTCACTTCACTTCACTTCACTTCACTTCACTCA 146909
QY 1876 CTGAGATCTCTTCTTATCTCTGTTATCTGCTAGTGAAGCTTGT 1921
Db 146910 CTGATAACCTTTCTTCCAGTTGATCAGTCACTGCTGCTGCTTCT 146955

RESULT 20
US-09-816-095-3/c
; Sequence 3, Application US/09816095
; Patent No. 6664084
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001147
; CURRENT APPLICATION NUMBER: US/09/816,095
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 99916
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1) - (99916)
; OTHER INFORMATION: n = A,T,C or G
US-09-816-095-3

Query Match 9.0%; Score 204.799; DB 1; Length 99916;
Best Local Similarity 50.1%; Pred. No. 5.5;
Matches 831; Conservative 0; Mismatches 722; Indels 106; Gaps 14;
QY 292 TCTATTTCTGATTTCTATCTTGGCTCATTTTAACTCAGTAGTGTGTTGGTTTCC 351
Db 40959 TTTATGCTTCTATTTCGTTATTACCAGTAGTCAATTCAGAGCAGGTTGTTCACTTCC 40900
QY 352 ATAAGTTTGAAGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 411
Db 40899 ATGCAGTTGTGCAGTTTGTAGTGAGT---TTCCTTAATCCTGAGTCTTAATTTGATTGCAT 40843
QY 412 GGTGTGATGATGACATAGAGTATTATTCATATGTTCTTTTATCTGTGAGAGATTGCTT 471

QY 542 TTGGTGAATAGTCTGTAAATATCT-CTAGGTCCACTTGGTTTATGACATCAGTTAGTCTC 600
Db 44800 TGGATGAATAGTCTGTATATCTGTAAATATCTTGTTCAGAGTATAGTTAAATC 44741
QY 601 CAGCATTTCTCTGTTTCTGTTTGTTCAGATGACCTAACTGTTCGAGAGATGGGGTAT 660
Db 44740 CATTCCTTCTCTGTTTCTGTTTGTTCAGATGACCTAACTGTTCGAGAGATGGGGTAT 44682
QY 661 TGAAGTACCCACATCTGTGTG-TGAGGTCAATATGATATTTAGCTGTAGCTGTCTC 719
Db 44681 TAAATCCCTTCACTATATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 44622
QY 720 GTTTTATGAACCTTGGGTGACATTTGTTTGTGATAGACATTAAGATTCGATGTCCT 779
Db 44621 GTTTTATGAACCTTGGGTGACATTTGTTTGTGATAGACATTAAGATTCGATGTCCT 44562
QY 780 CTGTGTGA-TTTTCTTTTGAGCTATGTAGTATTTTCCCAATCTC---ATCTGCTTA 835
Db 44561 CCTGTGAACAGGCTTTTACCATTATATAATGTCCCTCTTCTCTCTCTTTTAACTGC 44502
QY 836 GTTTTGGGTTTAACTATATAGTACATTAATAATGATCTATCGGCTGTCTCTTCTAGG 895
Db 44501 TGTACTTTAAAGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 44442
QY 896 GCCATTTGCTTGAATATCTTT-TCCATCTTTTACTCTAAGGTGATGCTATCCATCG- 953
Db 44441 TCCATTTGCTATGAATCTCTTCCCACTTTTACTTTTATGTTTATGTGAGTACTTATG 44382
QY 954 --TAGTGTCTTTTGGATGACAGTAGATGGAATCTGTTTTCATATCCATCTGT 1011
Db 44381 TACTAGGTGAGTCTCTCGAAGGACGACATAGTAGTGTGAGTCTCTATCCATCTGT 44322
QY 1012 TACCAGATATCTTTTCTAGAAATTAAGATCAATTCAGTCAATTCAGTGTGAGATATC 1071
Db 44321 GATTCGTATCTTTAAGTGGAGCATTTAGGCCATTTACATCCAAATGTAGTATGAAA 44262
QY 1072 AATGAGCAGTGTGTGATCTTGTGTT-----ATCTTCACCTGTGAA 1114
Db 44261 TGTGAGGATACCGTTGGCAATCATCATCTTTTGTTCCTGTCTACTTTGGGTTTTTT 44202
QY 1099 -----ATCTTCACCTGTGAA 1114
Db 44201 CCATAAAACCAAAAGAGCTAAGCTGTTCTCCCAACAGACCTTCAGTCTCCA 44142
QY 1115 GT 1145
Db 44141 GTGGGGGT 44082
QY 1146 TGT 1205
Db 44081 CACATTTTGTGGGGGT 44022
QY 1206 TCTCTCCCTCTTTTGTGATTTTGGCCT-----GGAATTTATTTATTTATTTCTT 1259
Db 44021 TTTAATCTGTATTTTATTTTGTAGATCCCTGTGTGATTTTGTGATTTTAAAGAGTTCATT 43962
QY 1260 GAATGTGGGTAACTTTTATGATGAATGATTTTCTCTACCTCTC-----TTTAGGT 1311
Db 43961 TTGATGTGTGTTCAGGATTTGTTTAAAGATTTAGAGTCTCTTTTAGCGGTCTTGTAGTG 43902
QY 1312 CTGCATTTGAGATAGATATCTTTTACATCTGATTTTATCTTGAATGCTTTCTTTCTC 1371
Db 43901 GTGACITGGTAATGGCAATTTCTCTCAGCATTTGTTGTCTGAAATGAGCTATCTTTTC 43842
QY 1372 CAACTATTTGACAGAAAGTTTCTTAAGTGCAGTAGTCTGGGCTGACATCTGTAGTCTC 1431
Db 43841 CTTTCATATAGATCTTTAGTTTTCACGTGATACAAATTTCTGGCTAATAATTTGTTTGTAG 43782
QY 1432 TTGGAGTCTAGACATCTGTGAGGGCTCTTCTATCTTTTGTAGTTTCTATTTGGAAG 1491
Db 43781 GAGGCCAGAG-----ATAGGGCTCCAATCCCTCTAGCTTTGAGGTTTCTGCTGAGAAA- 43727

QY 1492 TCAGGTGTAATCTTAATACATCTGCCTTATATATGTTAAATTTGGTCTTTTCCCTTCATC 1551
Db 43726 TCTGCTGTTAACTAATAGAGTTTTCCTTTGTAGGTTTCCCTGGT-GCTTCTGTCTCATATC 43668
QY 1552 TTTTAATATTTCTTTCTTTCTTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1611
Db 43667 TCTTAAGATTTCTTTCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 43608
QY 1612 GTTTCTTTTCCGTCCTCAATCTATTTTGGTGTGTTTGTATGCTTCTTGTACCTTGTATAGGCAT 1671
Db 43607 ATACCTTTTTCGATGAATATACCCAGGTGTTTTTGTGCTTCTTGTATTTGGATGTCTAG 43548
QY 1672 CTCTTTCTCAAGGTAGGAATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTCT 1731
Db 43547 ATTCTAGCAAGGCCGCGGGAAGTTTTTCTCTCAATATTTCCCTCAGATATGTTTTCGAAGCT 43488
QY 1732 TTTGACCTGCTTTCTT-----CCCTTCTCTTATCTCTTTGGTTTTTGGCATAGT 1780
Db 43487 TTTAAATTTCTCTTCTTCTCTCAGGGACACCATTAATCTTAGGTTGGTCAATTAACATA 43428
QY 1781 GTCTCTGGCTTCTCGATGTTTTTATGCTTGGATATTTTAGACTTAACAATTTCTTTTGAC 1840
Db 43427 ATATCAGACTTCTTGGAGTTTTTGTATATATTTCTTATTTCTTTTCTTTCTTTTCTT 43368
QY 1841 CAAAGGTATCCATTTCTCTATCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1900
Db 43367 TGAATGGGTAAATTTGAAGACTTTTGTCTTGGAGCTCTGAATTTCTTCTTCTTCTTCTT 43308
QY 1901 TATCTGTCTAGTGAGGCTT 1919
Db 43307 AATCTTATTTCTGACACTT 43289

RESULT 22
US-10-254-869-3/c
; Sequence 3, Application US/10254869
; Patent No. 6653117
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001160DIV
; CURRENT FILING DATE: 2002-09-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 148567
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)....(148567)
; OTHER INFORMATION: n = A,T,C or G
US-10-254-869-3

Query Match 8.8%; Score 200.5; DB 1; Length 148567;
Best Local Similarity 50.8%; Pred. No. 4;
Matches 894; Conservative 0; Mismatches 710; Indels 155; Gaps 19;
QY 305 TTCATCTTGGCTCATTTTAACTCAGTGTGAGTGTGTTTGGTTTCCATAAGTTTGTAAAG 364
Db 45036 TTTTGTGTTTGGACCAATGCCAATTCAGGAGCAGGTTAATTAATTTCCATGATTTGCATG 44977
QY 365 TTTTCTGTGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 424
Db 44976 GTTT---TGAGTTTCTTTTGGAGTTCATTTCCAGTTTATTTCCACTGTGCTCGAGAG 44920
QY 425 GACATAGAGTATTTCAATTTGCTTTTATCTGTGAGACTTGTCTTTGTTTGAATAT 484
Db 44919 AATGTTGA-TATAATGAATATTTCTAAATTTATTGAGGCTCATTAATATGCGCTATCAT 4861

[illegible]

RESULT 27

US-09-816-095-3

Sequence 3, Application US/09816095

; Patent No. 6664084

; GENERAL INFORMATION:

; APPLICANT: GAN, Wei

; TITLE OF INVENTION:

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

; TITLE OF INVENTION:
 ; TITLE OF INVENTION:

: FILE OF INVENTION:
: FILE REFERENCE: CLO

```

; FILE REFERENCE: CLO
; CURRENT APPLICATION

```

;
;
;

; CURRENT FILING DATE
: NUMBER OF SEC TO NO

```

; NUMBER OF SEQ ID NO
; SOFTWARE: FastSeq

```

; SOFTWARE: FASTSEQ I

Patent No. 6605432
GENERAL INFORMATION:
APPLICANT: Huang, Tim
TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR DETECTING DNA METHYLATION
FILE REFERENCE: UM01523
CURRENT APPLICATION NUMBER: US/09/497,855A
CURRENT FILING DATE: 2000-02-04
PRIORITY FILING DATE: 1999-02-18
PRIORITY FILING DATE: 1999-02-18
PRIORITY FILING DATE: 1999-02-05
NUMBER OF SEQ ID NOS: 54
SOFTWARE: PatentIn version 3.0
SEQ ID NO 40
LENGTH: 161652
TYPE: DNA
ORGANISM: Homo sapiens;
US-09-497-855A-40

Query Match 7.4%; Score 169; DB 1; Length 161652;
Best Local Similarity 48.3%; Pred. No. 5.1;
Matches 857; Conservative 0; Mismatches 800; Indels 116; Gaps 15;

QY 290 CTCTATTCTTGATTTCTATCTTGGCTCATTTTAACTCAGTAGTCAGTTGTTGTTT 349
DB 159703 CTACTAAACTTCCCTCTTACGCTTTACTCGTAGTCATTGAGAGTAAGTTT 159644

QY 350 CCATAAGTTTGTAAGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 409
DB 159643 CCATGTAGCTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 159584

QY 410 GTGTGTGTCAGATAGGACATAGATATTATTTCATTTGTTGTTGTTGTTGTTGTT 469
DB 159583 GTTGT- - - - -CTGAGAGACTGTTGTTGTTGTTGTTGTTGTTGTTGTT 159530

QY 470 TTTGTTTGAATATGATTAATCAATTTTGAG- -AGTTTCATAGGTCCTGACAGAGGT 527
DB 159529 TTTACTTCTAATTTTGGGCCAATTTTGAATAAGTACTATGTGAGGTTGAGAGAATGT 159470

QY 528 ACAGTCCTGTTGTTTGGTGAATAGTCTGTAATATCTCTAGTCCACTTGTGTTATGA 587
DB 159469 ATATCTGTTGATTTGGGTGGAGAGTTC- - - - -TGCTATTAGGTCGTCTGGTCAAGA 159415

QY 588 CATCAGTTAGCTCCAGCATTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 647
DB 159414 GCTCAGTTCAAGTCTCGAATGTCCTGTTGTTGTTGTTGTTGTTGTTGTTGTT 159355

QY 648 GAGATGGGTATTGAGTAGTCCACACTATCTGTGTGTGAG- -GTCAATATGTGATTTAGC 706
DB 159354 GACAGTGGGTGTTAAAGTCTCCACTATTATTGTTGGAGTATAAGTCTCTTTGTAGG 159295

QY 707 TGTAGCTGTCTGTTTATGAACCTTGGGTGACATTTGTTGTTGTTGTTGTTGTTGTT 766
DB 159294 TCTCTAAGACTCGTTTATGAATCTGAGTCTGTTATATGTTGGTGAATGATATTAGG 159235

QY 767 ATTCAATGTCCTCTTGGTGGATT- -TTCCTTGAATGCTATGATGTTTCTCCCAATCT 825
DB 159234 ATAGTTAGCTCTTCTGTTGATGATGATCACTTTAGCAATTAATGATGATGATGAT 159175

QY 826 CATCTGCTTAGTTTGGTTTAAAGTCTATT- - - - -AGTCAGATATAAAGTCTGTA 878
DB 159174 TTTAAGACCTTGTGTTTAAAGTCTGTTTATCAGAGACTAGGATTTCAATCCCTGCT 159115

QY 879 TCGGCTTGTCTTATGAGGCAATTTGCTTGAATATCTTT- -TCCATCTCTTTTATCTAAGG 937
DB 159114 TTTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 159055

QY 938 TGATGCTATCCATGGTAGGTTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 997
DB 159054 CTATGCTGTCTGACGCTGATGGGTCTCTGTAATCCGCACACTGATGGGTCTTGACTC 158995

QY 998 CATATCCATTCTGTTACCCAGTATCTTTTCTAGAGAAATTAAGATCATTTGATGATGA 1057

Db 158994 TTTATCAAAATTTGCCAGTCTGTCTTTT- - - - -AATGAAGCATTTAGCCCATTTA 158945

QY 1058 TGTGAGAAATTAATCAATGAGCAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1117

Db 158944 CATTTAAGTTAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 158885

QY 1118 TGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1177

Db 158884 AATTTGCTATTAGTTGATGAGTGTCTTCTTCAATGGGTCAAGGTCCTTTACATTTTGGTAT 158825

QY 1178 TCTGCTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1237

Db 158824 GTTTTGTAGTGGCTGCTACCGGTTTTTCTCTTCTTCC- - - - - 158790

QY 1238 TATTTATTATTTCATATTCTTTGTAATGTTGGTAAACATCTTTAGATTGAAGTTTCTCCT 1297

Db 158789 - - - - -ATATGATGCTCTCTTCAGG 158769

QY 1298 AGCCTCTTTAGTCTGCAATTTGAAGATAGATATTCTTTACATCTGATTTTATCTTAGAA 1357

Db 158768 AGCTCTGTTAAACACAGGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 158709

QY 1358 TGTCTTTCTTTCTCCAACTATTGTGACAGAAAGTTTCTTAAAGTCAGTAGTCTGSCCTG 1417

Db 158708 GGGTGTATTCTCTCTTCACTTAGGAGCTTAGTTGCTGATATGAATTTCTGGGTTG 158649

QY 1418 ACATCTGTAGTCTCTTGGAGTCTGTAGACATCTGTGAGGGCCTTCTTACATTTTGAAT 1477

Db 158648 AAAATCTTTCTTTAAAGAAATGTTGAATGTTGACCCCACTCTCTCTGGCTGTAGGCG 158589

QY 1478 TTCTATTGAAAGACAGGTAAATCTTAATACATCTGCTTTATATGTTTATTTGGTCTT 1537

Db 158588 TTCT-TCAGAGAGATCTGCTGTAGTCTGAT- -GGGCTTCCCTTGGGGGTAACTGACCT 158531

QY 1538 TTTTCCCTTGCATCTTTTAAATATTCTTTCTTTGTTCTATACCTTTTAGTGATTTGATTT 1597

Db 158530 TTCTCTCTGGCTGCCCTTAAACATTTTCTTCAATTCATCTAAGTGAATCTGACAAAT 158471

QY 1598 ATGCACTGTTGGAGTTCTTTTCCGTCCTCAATCTATTGTTGTTGTTGTTGTTGTTGTT 1657

Db 158470 ATGTG-TCCTGGATGCTCTTCTTGAAGGATATCTTGTGGTGTCTCTGTATTTCTCTGA 158412

QY 1658 ACCTTGATAGGCATCTTTTCTCAAGGTTAGAAATTTTCTTTTGGTTTCTTGTGAA 1717

Db 158411 AATTTGAATGTTGCTGCTGTTTCTAGTTTGGGAAGTTCTCTGATATATCTCTGAAGT 158352

QY 1718 ATATTTTCCCTGCTTTTGCACCTGCTCTTCTTCCCTTCTCT- - - - -CTATTCCT 1764

Db 158351 GTGTTTCCAACTTGGTTCCATTTCTCCCGTCACCTTTCAGGTACCAATCAACATAGCT 158292

QY 1765 TTTGGTTTTTGCATAGTCTGCTGGCTCTCTGATGTTTTATGCTGCTGATTTATTTAGCT 1824

Db 158291 TTAGTCTTTTCACTACTCTCATATTTCTTGAGGCTTTTTCATTCCTTTTCAATTTT 158232

QY 1825 TAACATTTTCTTTGACCAAGGTATCCATTTCTTCTATCTTCTTCTTCTTCTTCTTCTTCT 1884

Db 158231 TTCTCTATCTTGTCTTCTCATGCTTT- -AATTCATTAAGTTGCTCTTCAATCTCTGATATCC 158173

QY 1885 TCTCTCTATCTTCTTGTATTCTGTCAGTGTGCTCTCTGAGTTTCTGTTGTTGTTGTTGTT 1944

Db 158172 TTTCATCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 158115

QY 1945 TAAATTTTTCATTTCCAGATTTCTTCTGATTTGGGTTTTGTTTATTAATTTCTATTTCCAC 2004

Db 158114 TGCTGTGTTTTTGTAGCTCCATCAGTCAATTTATGTTGTTGTTGTTGTTGTTGTTGTTGTT 2037

QY 2005 TTTTCAAGTCTGTAAGATGTTTACTCATTTTCTCT 2037

Db 158054 TAGCAGTTCTCTGTAACCATTTTATAAAGGTTCTT 158022

RESULT 30

US-09-851-896-3

Db 24985 GTGATTTCTCTGGTGTATGTTTAAATTTCTGCTTTATTTTGT 25034

RESULT 31
US-09-803-671B-3/c
; Sequence 3, Application US/09803671B
; Patent No. 6582946
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001161
; CURRENT APPLICATION NUMBER: US/09/803,671B
; CURRENT FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 64467
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(64467)
; OTHER INFORMATION: n = A,T,C or G
US-09-803-671B-3

Query Match 6.4%; Score 145.4; DB 1; Length 64467;
Best Local Similarity 50.9%; Pred. No. 16;
Matches 595; Conservative 0; Mismatches 531; Indels 42; Gaps 11;

QY 292 TCTATTTCTTGTATTTCTATCTTGGCTCAATTTTAACTCAGTAGTGAGTTGTTGTTCC 351
Db 54901 TTTTCTTCTTAAATTTCTTCTTGGCCAGTAGCCATTGAGAACTTTGTTAGTTTCC 54842

QY 352 ATAAAGTTTGTAAATTTCTGTTGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 411
Db 54841 ATGTAATTTATACAGTTTCCAAAGTTCTCTTGTATTTGTTGTTGTTGTTGTTGTTGTT 54782

QY 412 GGTGTCAGATAGGACATAGATATTTTCAATTTGTTTATCTGTCGAGACTTGTCT 471
Db 54781 GG---CTGAAGAATATCTGTAATGTTTGAATTTTAAATTTTGAATTTGTTT 54726

QY 472 TGTGTTGAAATGTAATTTCAATTTTGGAGAGTTTCAATAGGTTGTCGACAAAGGTACAG 531
Db 54725 TGCAGCCTAATGATATGGCTATCTCTGAAATGTTTCCATGTTTAAAGGAACTTTAT 54666

QY 532 TCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 591
Db 54665 CTGCAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 54606

QY 592 AGTTAGCTCCAGCAATTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTT 651
Db 54605 AATTTAAATCCAGGTTTCTCT-ATTGATGTTTGTGTTAGATGATCTGTTTAAATGTTGAGA 54547

QY 652 ATGGGTTATTTGAGTACCCATCTCTGTTGTTGAG-GTCAATATGTTGTTTCTGTTGTT 710
Db 54546 GTGGGTTGTTAAAGTCCCATCTATGATTTGTTTATGTTGTTTATCTCTCTTCAGATCTA 54487

QY 711 GCTGTTGTTGTTTATGAACTTTGGGTGACATTTGTTTGGTGCATPAGACATTTAAGAATTG 770
Db 54486 ATAAATTTGCTTCATATACCTGGGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 54334

QY 771 CAATGTTCTTTGTTGTTGTTTCTCT-TTGATGCTTATGTTGTTTCTTCCCAATCTCATC 829
Db 54433 TTACATCTCTGTTGTTGTTTCTCTTCTTATCATTTATGTTGTTTCTCTCTCTCTCTA 54374

QY 830 TGTGTTGTTTGGTTTAAAGTCTTATGTTGTTTAAATGTTTAAATGTTTAAATGTTTAAATGTT 889
Db 54373 TGT---TTTGACATAAGAGTGGTTTATCTGATATGATGTTTATCTCTCTCTCTCTA 54317

QY 890 CTTAGGGCAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 948

Db 24985 GTGATTTCTCTGGTGTATGTTTAAATTTCTGCTTTATTTTGT 25034

RESULT 31
US-09-803-671B-3/c
; Sequence 3, Application US/09803671B
; Patent No. 6582946
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001161
; CURRENT APPLICATION NUMBER: US/09/803,671B
; CURRENT FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 64467
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(64467)
; OTHER INFORMATION: n = A,T,C or G
US-09-803-671B-3

Query Match 6.4%; Score 145.4; DB 1; Length 64467;
Best Local Similarity 50.9%; Pred. No. 16;
Matches 595; Conservative 0; Mismatches 531; Indels 42; Gaps 11;

QY 292 TCTATTTCTTGTATTTCTATCTTGGCTCAATTTTAACTCAGTAGTGAGTTGTTGTTCC 351
Db 54901 TTTTCTTCTTAAATTTCTTCTTGGCCAGTAGCCATTGAGAACTTTGTTAGTTTCC 54842

QY 352 ATAAAGTTTGTAAATTTCTGTTGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 411
Db 54841 ATGTAATTTATACAGTTTCCAAAGTTCTCTTGTATTTGTTGTTGTTGTTGTTGTTGTT 54782

QY 412 GGTGTCAGATAGGACATAGATATTTTCAATTTGTTTATCTGTCGAGACTTGTCT 471
Db 54781 GG---CTGAAGAATATCTGTAATGTTTGAATTTTAAATTTTGAATTTGTTT 54726

QY 472 TGTGTTGAAATGTAATTTCAATTTTGGAGAGTTTCAATAGGTTGTCGACAAAGGTACAG 531
Db 54725 TGCAGCCTAATGATATGGCTATCTCTGAAATGTTTCCATGTTTAAAGGAACTTTAT 54666

QY 532 TCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 591
Db 54665 CTGCAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 54606

QY 592 AGTTAGCTCCAGCAATTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTT 651
Db 54605 AATTTAAATCCAGGTTTCTCT-ATTGATGTTTGTGTTAGATGATCTGTTTAAATGTTGAGA 54547

QY 652 ATGGGTTATTTGAGTACCCATCTCTGTTGTTGAG-GTCAATATGTTGTTTCTGTTGTT 710
Db 54546 GTGGGTTGTTAAAGTCCCATCTATGATTTGTTTATGTTGTTTATCTCTCTTCAGATCTA 54487

QY 711 GCTGTTGTTGTTTATGAACTTTGGGTGACATTTGTTTGGTGCATPAGACATTTAAGAATTG 770
Db 54486 ATAAATTTGCTTCATATACCTGGGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 54334

QY 771 CAATGTTCTTTGTTGTTGTTTCTCT-TTGATGCTTATGTTGTTTCTTCCCAATCTCATC 829
Db 54433 TTACATCTCTGTTGTTGTTTCTCTTCTTATCATTTATGTTGTTTCTCTCTCTCTCTA 54374

QY 830 TGTGTTGTTTGGTTTAAAGTCTTATGTTGTTTAAATGTTTAAATGTTTAAATGTTTAAATGTT 889
Db 54373 TGT---TTTGACATAAGAGTGGTTTATCTGATATGATGTTTATCTCTCTCTCTCTA 54317

QY 890 CTTAGGGCAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 948

Db 54316 TTGGTTTCTATTTCATGAACATACTCTTTTCCATCCTTTTCCATCTTACGTCTATATGTGTC 54257
Qy 949 CAT---GGTAGGTGTCTTTTGGATGGACAGTAGGATGATCTTGTGTTTTCATATCCA 1005
Db 54256 TTTACAGGTAAAGTGAGTTCTTTGAAAGCAGTATATTTTGGGTCACTTTTAAAAATC- 54198
Qy 1006 TTCTGTTTACCAGTATCTTTTCTAGAGAAATTAAGATCATTTAGTCATTGATGTTGAGA 1065
Db 54197 -----TATTCAGCCATTAAATTTGGAATTTAATCCATTTATGTTTATTCCTGTCA 54144
Qy 1066 AATTATCAATGAGCAGTGTGTTGTAATCTTGTATCTTTGCACTTTGGAAGTGTGTGTG 1125
Db 54143 TTTTGTGATGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 54084
Qy 1126 TG 1185
Db 54083 TCTCTGTTTATTAATCATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 54031
Qy 1186 TG 1245
Db 54030 AGTCCTTCTCTACCTCTTTGT 53971
Qy 1246 ATTATCATTTTCTTGATGTTGGTAAACATCTTTAGATTGAATTTTCTCTAGCCTTCT 1305
Db 53970 GTTTTCAATGATGATGATATGACCTTTTCACTTCCAGGTGAGGACTTTCTTAAGCAATTC 53911
Qy 1306 TTAGGTCTGATTTGAAGATAGATATTTTATCATCTGATTTTATCTTAGAATGTCTTTC 1365
Db 53910 TTGTAGGCCACATTAATGGTGTGATGAGTCTCTCAGTTTTCATGCTGTGTGAAGACTTTC 53851
Qy 1366 TTTCTCCAACTATTGTGACAGAAAGTTTCTTAAGTCCAGTGTCTGGCCTGACATCTGT 1425
Db 53850 T-----CCATTTTGAAGGATAACTTTGCTGAGTATAGCATTTCTTGGCTGACAGTTT 53798
Qy 1426 AGTCTCTTGGAGTCTGTAGCACATCTGT 1453
Db 53797 GTTTGTTTATGCCCTTCAAAACTTT 53770

RESULT 32

US-09-393-634-8
; Sequence 8, Application US/09393634
; Patent No. 6556910
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Ryba, Nick
; APPLICANT: Mueller, Ken
; APPLICANT: Hoon, Mark
; APPLICANT: The Regents of the University of California
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: SF, a No. 6558910el Family of Taste Receptors
; FILE REFERENCE: 02307E-098000US
; CURRENT APPLICATION NUMBER: US/09/393,634
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 8
; LENGTH: 1540
; TYPE: DNA
; ORGANISM: Rattus sp.
; FEATURE:
; OTHER INFORMATION: rat GR04 sequence approximately 1100 bp 5' to SEQ
; OTHER INFORMATION: ID NO:92
US-09-393-634-8

Query Match 6.3%; Score 142.8; DB 1; Length 1540;
Best Local Similarity 67.1%; Pred. No. 6.3e+02;
Matches 286; Conservative 0; Mismatches 132; Indels 8; Gaps 6;
Qy 589 ATCAGTTAGTCCAGCATTTCTCTGTTTCGTTTCTGTCAGATGACCTAACTGTTGGAG 648

Db 1115 ACTGTTAGTTCTCTGTGTCCTCTGTTAG-TTTCTGTTTCCATGATCTGTCATGATG 1173
Qy 649 AGAATGGGGTATTGAAGTAGCCCACTATCTGTGTGAGGTC-AATATGTGATTTTAGCT 707
Db 1174 AGCGTGGGGTGTGAAATCTCCGACTATTGTTGTGTGAGATGAAATGTGTCTTTGAGCT 1233
Qy 708 GTAGCTGTGCTGTTTATGAATTTGGGTGACATTTGTTGGTGCATAGACATTAAGAA 767
Db 1234 TTAGTAAGATTTCTTTTGTGAATGTAGGTGCTTTTGGCAATTTGGTGCATAGATATTAA 1293
Qy 768 TTGCAATGTCCTTTGGTGGGA-TTTTCCCTTTGATGCCCTATGTAGTATTTCTCCCAATCTC 826
Db 1294 TTGAGAGTTTCACTTTGGTGGATTTTCCCTTTGATCAATATGAAGTGTCTCTTCTATCTT 1353
Qy 827 ATCTGCTAGTTTGGGTTTAAAGTCTA-TTAGTCAATATTAAGATGACTGTATCGGCTT 885
Db 1354 TTTTGATGACTTTTGAAGACGTCATTTTATGGAATTAGATTGGAACACTCAAGATT 1413
Qy 886 GGTTCCTTAGGGCCATTTGCTTAG-AATATCTTTTCCATCCTTTTACTCTAAGGTGATGTC 944
Db 1414 GCTTCTGAGGTCAITTCCTTGGAAAGTGTTTTTCAGCCATTTACTCTGAGGTAGTGT 1473
Qy 945 TATCCATG---GTAGGTGTCTTTTGGATGACAGCAGTAGGATGATCTTGTGTTTCATA 1001
Db 1474 TGTCTTTGTCTCTGAGGTGTGTTTCTCTCAITTCAGCAAAATGCTGGTCTCTTTACATA 1533
Qy 1002 TCCATT 1007
Db 1534 TCCAGT 1539

RESULT 33

US-09-220-132-168/c
; Sequence 168, Application US/09220132
; Patent No. 6506807
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
; OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220,132
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068,821
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 168
; LENGTH: 9573
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-220-132-168

Query Match 6.0%; Score 136.8; DB 1; Length 9573;
Best Local Similarity 57.6%; Pred. No. 1.1e+02;
Matches 333; Conservative 0; Mismatches 237; Indels 8; Gaps 5;
Qy 292 TCTATTTCTGATTTCTATCTTGGTCTATTTTAACTCAGTAGTGAGTTGTTGTTTCC 351
Db 4147 TTTCTCTTCTGATTTCAITTTAACTCCAAAGGTCTTCAGGAGCAGATATTATTTTC 4088
Qy 352 ATAAATTTGTAAGTTTCTGTTGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 411
Db 4087 ATGTAATTTGTATAGTTTAAAGGTT---TCTTTTGGAGTTAAATTTCCAGTTTATTCAT 4031
Qy 412 GGTGTGCAGATAGGACATAGAGTATATTCAATTTGTTCTTTTATCTGTCGAGACTTGCT 471
Db 4030 TGTGTCTGAAAGGTACTTGACATAATTTCAATTTCTTAAATTTGTTGAGACTTGTT 3971
Qy 472 TCTTTTGAATATGATTCAATTTTGGAGA--GTTTCATAGGGTCTGACAGAGAGTAC 529

Db 3970 TGTGGCCTATCTTATGATTTATCTTGGAGAACGTTCTATGCTGCTGATTAATAATGGTAT 3911
Qy 530 AGTCT-TTGCTGTTTGGTGAATAGTCTTAATAATCTCTAGTCCACATTTGGTTATGAC 588
Db 3910 ATTCGAACTTGGGTAGAAATGTTAGGTAAATAATCTGCTAAATCTATTTGTTCTAGGA 3851
Qy 589 ATCAGTTAGCTCCAGCAATTTCTCTGTTTCGTTTTTTTTCAGATGACCTAACTGTTGGAG 648
Db 3950 TATTAATTAAGTCAATTTTCTTTATTTGACTTTCTGCTTTGATGCTCTAGTCTG 3791
Qy 649 AGAATGGGTATTAAGTAGCCCACTATCTGTGT-GTGAGGTCAATATGATGATTTAGCT 707
Db 3790 TCCATGGAGTATGAAGTCCCACTACTATTATGTTGGTGGCATCTATCCCAATTTCTTAGGT 3731
Qy 708 GTAGCTGTCTGTTTATGAACCTTGGGTGACATTTGTTGGTGCATACACATTAAGAA 767
Db 3730 CTAGTAATTAATGTTTATTAATTTGGAGCTCCAGTATTAGTGCATATATTTAGGA 3671
Qy 768 TTGAATGTCTCTGTGTGGA-TTTTCTTTGATGCTATGTAGTATCTTCCCAATCTC 826
Db 3670 TTGTGATATTTGTTCTGT-TGGACCATCTTTATCATTAATAAATTTCCCTTTCTTGTCT 3611
Qy 827 ATCTGCTTAGTTTGGGTTTAAGTCTATTAGTCAGATA 864
Db 3610 TTATAACTGTTGTGCTTTAAAGCTTGTCTGATATAAA 3573

RESULT 34
US-09-146-053-4
; Sequence 4, Application US/09146053A
; Patent No. 6399349
; GENERAL INFORMATION:
; APPLICANT: Ryan, James W.
; APPLICANT: Sprinkle, Terry Joe Curtis
; APPLICANT: Venema, Richard C.
; TITLE OF INVENTION: Human Aminopeptidase P Gene
; FILE REFERENCE: MGI103
; CURRENT APPLICATION NUMBER: US/09/146,053A
; CURRENT FILING DATE: 1998-09-02
; EARLIER APPLICATION NUMBER: 60/057,854
; EARLIER FILING DATE: 1997-09-02
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 50000
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-146-053-4

Query Match 5.5%; Score 125.5; DB 1; Length 50000;
Best Local Similarity 54.2%; Pred. No. 25;
Matches 429; Conservative 0; Mismatches 325; Indels 37; Gaps 9;

Qy 268 CATTCCCTCTCTCTCCAAACACTTCTATTTCTTGATTTCTATCTTGGCTCATTTTAAAC 327
Db 28221 CATTCAATTAATAATACATTTCTAATTTCTTTTGACTTCTCTTTGACTATGGAATAT 28280
Qy 328 TCAGTAGTGAAGTTTGGTTTCATTAAGTTTGTGAAGTTTCTGTTTCTGTTGTTGT 387
Db 28281 TTAGAAGTGTGTTATTTAGTTTCCAGATATCTAGGCCATTTTCCAGAG---ATCTTCTGT 28337
Qy 388 TGTGTTATCTAGATTTAAAGCTGTGGTGCAGATAGACATAGATTT-----437
Db 28338 TATTGATTTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 28397
Qy 438 ATTTCAATGTCTTTATCTCTCGAGACTTCTCTTTGTTTGGTAAATATGATTAATTTT 497
Db 28398 ACTTGAATTTCTTTAAATTTCTTGAATTTCTTTTATGGCCCAAGATATGCTCTATCT 28457
Qy 498 G--AGAGTTTCATAGGGTGTGCAAGAGAGTACAGT-CTTTGTTGTTTCTGTTGAATAGT 554
Db 28458 GTAAATGTCTGTATGCTTTTGGAAAGAAAGTATATTTCTCTCTCTTGTGGTGGAGTCT 28517

Qy 555 CTGTAATATCTCTAGGTCCACTTGGTTTATGACATCAGTTAGTCCAGCAATTTCTCTGT 614
Db 28518 CTATAAATATCAATTTAGGTCAAGTTGGTTGATAGTTTATTCATATCTCTATATCTTG 28577
Qy 615 TTCGTTTTTGTGAGATGACCTAACTGTTGGAGAGATGGGTATTCAGATGACCCACT 674
Db 28578 CTGATCTCTCTTCTTCTTATTTCTATCAATTAATGAAGATGGGCTTTGAAA-----28628
Qy 675 ATCTCTGTGTGAGGTCATATGTCATTTTAGCTGTAGCTGTGCTTGTGTTTATGAACTTGG 734
Db 28629 -----TCTGCAATTTCTATTTCTACTTGAAGTCTGTGAGTTTACTTCAGGTATTTG 28683
Qy 735 GTGCAATTTGTTGGTGCATAGACATTAAGAATTCATGCTCTCTTGTGGATTTCTTTC 793
Db 28684 AAGCTCTGTATTTGGGTACATAAGCGTTTGAATAATGTTATGCTCCCTTGATGAATGACC 28743
Qy 794 CTTTGATGCCATGTPAGTATTTCTCCCAATCTCATCTGCTTAGTTTGGGTTTAAAGTCTA 853
Db 28744 CTTTCATCACATGAATAATTAATCTTTTACTCTGGAATTTATTTTGTCTTTGAAATCTA 28803
Qy 854 -TTAGTCAGATATTAAATAGACTGTATCGGCTTCTCTTAGGGCCATTTGCTTAGAATA 912
Db 28804 CTTTATCTCATATTAATATGACATTTTCAGATTTCTTTTGAATTAGTGTAGCATGTATA 28863
Qy 913 --TCTTTTCCATCTTTTACTCTTAAGGTGATGCTATCCATGTTAGGTGTTGTTTTTGG 970
Db 28864 ACTTTTTTCCATACCTTTTAACTATTTCTAACT---TTGGAGTTTAAATGTTTTCTTAT 28920
Qy 971 ATGCAGCAGTAGGATGATCTTTGTTTTCATATCCATCTCTTACCCAGTATCTTTTCTA 1030
Db 28921 AGCGAGTACATAGTTGGCTGTGTTGTTTATATCAATCTCGAAATCTCAATCTTTATTT 28980
Qy 1031 GAGAAATTAAG 1041
Db 28981 GGGTATTTAAG 28991

RESULT 35
US-09-525-160B-9/C
; Sequence 9, Application US/09525160B
; Patent No. 6569681
; GENERAL INFORMATION:
; APPLICANT: Ivanov, Evgenii
; TITLE OF INVENTION: METHODS OF IMPROVING HOMOLOGOUS RECOMBINATION
; FILE REFERENCE: 10278/016001
; CURRENT APPLICATION NUMBER: US/09/525,160B
; CURRENT FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 3033
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-525-160B-9

Query Match 5.1%; Score 117.1; DB 1; Length 3033;
Best Local Similarity 58.5%; Pred. No. 4.2e+02;
Matches 340; Conservative 0; Mismatches 224; Indels 17; Gaps 8;

Qy 379 TCTTGTGTTGTTGTTATCTAGATTTAAGCTGTGTGCTCAGATAGGACATAGACTATTA 438
Db 2079 TGTATTTATTTGGGACTTCTTTTCTTTCTAGTCTGTGTGAGAGAAACGTTTGATATTA 2020
Qy 439 TTTCAATTTGCTTTTATCTGTCGAGACTTGCTTTGTTTGAATATGTTATTCATTTTGG 498
Db 2019 TTTCACTTTTCTTAAATTTTAAATTTTGTGACCTAGCATTTGGTATATCTCTTG 1960
Qy 499 AGAGTTT--CATAGGGTCTGCAAGAAGGTACAGTCTTTGTGTTTGG-TGAATAGTC 555
Db 1959 AGAATTTCCATGCTCTGAGGAGAGAAATGTGTTATTTACAGCTGTTGGATCAATGTC 1900
Qy 556 TGTAAATATCT-CTAGGTCCACTTGGTTTATGACATCAGTTAGTCCAGCAATTTCTCTGT 614

Db 1899 TGTAAATATCTACCTAGGTCACCTTTGACCTATGGTGCAGATGAATTTGATGTTTG----- 1845
Qy 615 TTCTGTTTTTTGTTGAGATGACCTAACTGTTGGAGAGAAATGGGTATTGAAGTAGCCCACT 674
Db 1844 TTAATTTTCTGCTGGATGATCTGTCCTCAATGCTGAAAGTGGGTGTTGAAGTCTCCAGTT 1785
Qy 675 ATCTGTGTGAG-----GTCAATATGTGATTTTACGTGTAGCTGTGCTGTTGTTTAA 729
Db 1784 ATTATGCTATGGGGTCTGCTATCTCTCTTTGGCTCTAAATATTTGCTTTTATATA 1725
Qy 730 CTTCGGGTGACATTTGTTTGGTGCATAGACATTAAGAAATGCAATGTCCTCTTTGGTGGAT 789
Db 1724 CTGAGTACTCCAGTATGGTGNATATATTTAGAAATTTGTAATGCTCTTTGCTAAAT 1665
Qy 790 TTTCCT--TTGATGCTATGAGTATTTCCCAATCTCATCTGCTTATTTGGGTTAA 848
Db 1664 GGACCCCTTTATCACTATAATGACCTTTCTTCTCTCTTTTCATAGTTTTTGTCTTGCA 1605
Qy 849 GTCTA--TTGATGCTATGAGTATTTCCCAATCTCATCTGCTTATTTGGGCTTATC 948
Db 1544 GAATAT--TTTTCCATCCCTTTACCTTCAGTCTATATGTGC 1505
Qy 908 GAATATCTTTTCCATCCCTTTACTCTAAGGTGATGCTATC 948
Db 1604 GTCTATTTTGTCTGATATAATGAGTCTTCTGCTCTTTTGTGTTTCCATAGGCATA 1545
Qy 908 GAATATCTTTTCCATCCCTTTACTCTAAGGTGATGCTATC 948
Db 1544 GAATAT--TTTTCCATCCCTTTACCTTCAGTCTATATGTGC 1505
RESULT 36
US-09-525-160B-4/c
; Sequence 4, Application US/09525160B
; Patent No. 6569681
; GENERAL INFORMATION:
; APPLICANT: Ivanov, Evgenii
; TITLE OF INVENTION: METHODS OF IMPROVING HOMOLOGOUS RECOMBINATION
; FILE REFERENCE: 10278/016001
; CURRENT APPLICATION NUMBER: US/09/525,160B
; CURRENT FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 3213
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-525-160B-4
Query Match 5.1%; Score 117.1; DB 1; Length 3213;
Best Local Similarity 58.5%; Pred. No. 4e+02;
Matches 340; Conservative 0; Mismatches 224; Indels 17; Gaps 8;
Qy 379 TGTGTTGTTGTTGTTTATCTAGATTTAAAGCTGTGGTGCAGATAGGACATAGATTA 438
Db 2079 TGTATTTATTTGGACTCTTTTCTTCTAGTCTGTGTCGAGAGAAAACGTTGATATTA 2020
Qy 439 TTTCAAATGCTTTTATCTGTCGAGACTTGCCTTTGTTTGGAAATATGATTTCAATTTGG 498
Db 2019 TTTCACCTTTTCTTAAATTTTAAAGATTTGTTTGTGACCTAGCATTTGGTATATCCTTG 1960
Qy 499 AGAGTTT--CATAGGGTGTGACAAGAGGTACAGTCTTTTGTGTTTGG--TGAAATAGTC 555
Db 1959 AGAATTAATCAATGCTGAGGAGAGAAATGTGATTTCTACAGCTGTTGGATCAATGTTTC 1900
Qy 556 TGTAATATATCT--CTAGTCCACTGTTGTTTATGACATCAGTTAGCTCCAGCATTTCTCTGT 614
Db 1899 TGTAATATATCTAGGTCCATTTGACCTATGGTGCAGATGAATTTGATGTTG----- 1845
Qy 615 TTCTGTTTTTTGTTGAGATGACCTAACTGTTGGAGAGAAATGGGTATTGAAGTAGCCCACT 674
Db 1844 TTAATTTTCTGCTGGATGATCTGTCCTCAATGCTGGAAGTGGGTGTTGAAGTCTCCAGTT 1785
Qy 675 ATCTGTGTGAG-----GTCAATATGTGATTTTGTAGCTGTAGCTGTGCTTTTATGAA 729
Db 1784 ATTATGATGGGGTCTGCTATCTCTCTTTGGCTCTAAATATTTGCTTTTATATA 1725

Qy 730 CTTCGGGTGACATTTGTTGTTGTCATAGACATTAAGAAATGCAATGCTCTCTTGTGAT 789
Db 1724 CTGAGTACTCCAGTATTGGGTGAATATATATTAGAATTTGTAATGCTCTTGTCTAAAT 1665
Qy 790 TTTCCT--TTGATGCTATGAGTATTTCCCAATCTCATCTGCTTATTTGGGTTAA 848
Db 1664 GGACCCCTTTATCACTATAATGACCTTTCTTCTCTTTTTCAPAGTTTTTGTCTTGCA 1605
Qy 849 GTCTA--TTAGTGCAGATATTAATAATGACTGTATCGGCTTGTCTTTAGGCCCATTTGCTTA 907
Db 1604 GTCTATTTTGTCTGATATAATGAGTCTTCTGCTCTTTTGTGTTTCCATAGGCATA 1545
Qy 908 GAATATCTTTTCCATCCCTTTACTCTAAGGTGATGCTATC 948
Db 1544 GAATAT--TTTTCCATCCCTTTACCTTCAGTCTATATGTGC 1505
RESULT 37
US-09-818-512-3/c
; Sequence 3, Application US/09818512
; Patent No. 6537780
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO01192
; CURRENT APPLICATION NUMBER: US/09/818,512
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 116592
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)--(116592)
; OTHER INFORMATION: n = A,T,C or G
US-09-818-512-3
Query Match 5.1%; Score 116; DB 1; Length 116592;
Best Local Similarity 51.8%; Pred. No. 12;
Matches 565; Conservative 0; Mismatches 465; Indels 60; Gaps 14;
Qy 212 ACTGCTCTCTCTCCCTTCTCTAACACTTCTGCGCCAGGTAGGGGCACCTACCGCAT 271
Db 100508 ATGTCTATACACTTCCCTCTTAGAACCCGCTTTTGTCTGTATCCCATAGGTTTTGGTATTTT 100449
Qy 272 CCCTCTCTCTTCCAAACACTTCTATTTCTGATTTCTATCTTGGCTCATTTTAACTCAG 331
Db 100448 GTTTCATTTTCATTTGCTCAAGAAATTTTAAATTTTACITTAATTTCTTCTATTGTT 100389
Qy 332 TAGTGAGTTGTTGGTTTCCATAAGTTTGTGAAGTTTTCGTGTTTCTGTTGTTGTTGTT 391
Db 100388 CCATTTGTTTAAATTTCCATTAATTTGTACTGTTTCCAATGTTCC--TCTGATTATA 100332
Qy 392 GTATCTAGATTTAAGCTGTGGTGTGATAGGACATAGAGTATTTATTTCAATTTGCTT 451
Db 100331 GAITTTCTAGTTTATTTCCACGTGTGACAAGATA-----AATATAATTTCAATTTTAA 100281
Qy 452 TTATCTGTGAGACTGTTGTTTGTGAAATATGATPANTCAATTTTGGAGA--GTTTCATA 509
Db 100280 AAAATTTGTAAACCTTTGTTGTGACCTTAATGTATGATCAATCTCTGGAGAAATTTCCCATG 100221
Qy 510 GGGTCTGCACAGAGAGGTACAGTCTTTGTTGTTTGG--TGAAATAGTCTGTAATATCT-C 567
Db 100220 TGTCTGTGAGAAAAATGTTATTTCTGCAACTGTTGGATGGAAATGTTCTGTAATATCTGT 100161
Qy 568 TAGGTCCACTTGGTTTATGACATCACTAGTTAGCTCCAGCATTTCTCTGTTTCTGTTTCTGTT 627
Db 100160 TAGGTCCACTTGGTCCAGAGACAGTTTAAATGCTCTACGTTTCTTGTGTTTCTTCTGCTCT 100101

628 GAGATGACCTAACTGTTGGAGAGAAAGGGTATTGAAGTAGACCCACATATCTGTGT-GTGA 686
Qy
Db 100100 GCATGATCTGTTCAATTTGAAA-----GTGGGATGTTCCCTAGTATTACTGTATTGC 100048
Qy 687 GGTCAATATGTAATTTAGCTGTAGCTGCTGCTTTTATGAAGTGGGAGACATTTGTGT 746
Db 100047 AGTCTATCTCTCTTTAGGTCTATTAAATATTGCTTTTATATATTGGGTGCTCTGGTGT 99988
Qy 747 TTGGTGCATAGACATAAGAAATTCGAATTCCTCTGCTGCTGGA-TTTTTCTTTGATGCCCTA 805
Db 99987 TGGGTGATATATATTTACCATTTGTTATATCTCTCTGCTGCTGACCCCTTTGTCGTTA 99928
Qy 806 TGTAGTATCTTCCCAATCTCACTGCTAGTTTGGGTTTAACTGCTATTAGTCAGATAT 865
Db 99927 TATATGCGCTTGTCTTTTAAAC---ATCTTGATTTAAGTTTATTATTATGTCATAA 99871
Qy 866 TAAATGACTGATATCGGCTGCTCTTTAGGGCCATTTGCTTAGAATATC-TTTTCCATCC 924
Db 99870 AACTATAGGCTCTCTGCTCTTTTGGTTTCCATTTGCAATTAATTAATCTTTTCCATC 99811
Qy 925 TTTTACTCTAAGGTGATGCTATCCAT---GGTAGGTGCTCTTTTGGAGTCAGCAGTA 981
Db 99810 CTTCACTTTAGTCTGTGCTATCTTTTACAGATGAAGTGAGTTCTTTGTAGCCAGCATAT 99751
Qy 982 GGATGGATCTGTTTTCATATCCATCTGTTACCCAGTATCTTTTCTAGAGAAATTAAG 1041
Db 99750 GGATGGGTCTGTTTAAATCCATCTGCCACTTT-----AATGAA 99707
Qy 1042 ATCATTTAGTCAATGATGTTGAGAAATTAATCAATGAGCAGTGTTTTGGATCTTGTATC 1101
Db 99706 GAATTTAATCCATTATATTCAAAGTTATTACTGACAGGTAAGGGTTTACTAGCCATA 99647
Qy 1102 TTGACATTTGCAAGTG 1161
Db 99646 CTGTTCCT-----TGTTTTGGTGTGCTTTGTAGATCTTTTCTCTATTTTCC 99599
Qy 1162 TG 1221
Db 99598 TCTCTTACTGCTTCTCTCTAGTAAAGTATTTTCTAGCAGGTGCTTTGATTTTG 99539
Qy 1222 ATTTTGGCCCGAATTAATTAATTAATTAATTTTCTTGAATGGGTAACATCTTTAGA 1281
Db 99538 CTATTTATTTTACTGTAATTAATTAATTAATTTTGTCTTTTGTGTTACTGTAGGCTTACA 99479
Qy 1282 TTGAAGTTT 1291
Db 99478 AAAAAATCTT 99469

RESULT 38

US-09-734-673-3
; Sequence 3, Application US/09734673
; Patent No. 6410294
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001020
; CURRENT APPLICATION NUMBER: US/09/734, 673
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 38564
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(38564)
; OTHER INFORMATION: n = A,T,C or G
US-09-734-673-3

Query Match 5.1%; Score 115.3; DB 1; Length 38564;
Best Local Similarity 59.1%; Pred. No. 36;
Matches 298; Conservative 0; Mismatches 197; Indels 9; Gaps 6;
Qy 298 TCTTCATTTCTATCTTTGGCTCATTTTAACTCAGTAGTGAGTTGTTGGTTCCATAAGT 357
Db 22517 TCTTTATTTCTATTGTTAAACCCAAATTTATACAGGAGCAGATTTTAAATTCATGTAT 22576
Qy 358 TTGTAAGTTTCTGTTGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 417
Db 22577 TTGTATAGTTTGGAAATTCCTTTTGGAGTTGGTT---TCTAGTTTTCCTACTGTGT 22633
Qy 418 CAGATAGACATAGAGTATTTTCAATTTGCTTTTATCTCTCGAGACTTCTGTTGTTT 477
Db 22634 CTGAGAAATGCTTGATATGATTTTGAATTTTAAAT- TATTGATCTTGTGTTGAGC 22692
Qy 478 GAAATATGTTATTTCAATTTTGA--GAGTTTCATAGGGTCTGCAAGAGGTACAGTCT- 534
Db 22693 CAATCGTATGCTTATGTTGGAATATATTCATGCTGATGAGAGAAAGTATATCTA 22752
Qy 535 TTGTTTGGTGAATAGTCTGTAATATCTCTAGGTCCACTTGGTTTATGACATCAGT 594
Db 22753 TAGTTTGGGTAGGATGCTCTGTAATATCTGTTAGGTCTACTGTTTGTAGATATTGT 22812
Qy 595 TAGTCTCCAGCAATTTCTCTGTTTCTGTTTCTGTTGATGACCTAACTGTTGGAGAGATG 654
Db 22813 TTAAGTCTGTTCTTTTG-TTGACTTCTGCTCGATGATCTGCTAGTCTGTCTAGTG 22871
Qy 655 GGGTATGAAGTAGCCCACTATCTGTGTGTG-TGAGTCAATATGATGATTTTGTAGTGTAGCT 713
Db 22872 GGGTGTGAAGTATCCCACTATTATTGGGTTGCTCAITATCTCTTCCCTAGGTCTAGCA 22931
Qy 714 GTGCTGTTTATGAACCTTGGTGACATTTGTTGGTGCATAGACATTAAGAATTGCAA 773
Db 22932 GTAATGTTTATGAATTTTGAGCTCCATAGTTAGGTACATATATTTGGATTATAA 22991
Qy 774 TGTCTCTTGGTGGATTTTCTTTT 797
Db 22992 TATTCTGTGACAGATTAATCCTT 23015

RESULT 39

US-09-010-398-14
; Sequence 14, Application US/09010398
; Patent No. 5972615
; GENERAL INFORMATION:
; APPLICANT: An. Gang
; TITLE OF INVENTION: BIOMARKERS AND TARGETS FOR DIAGNOSIS, DISEASE
; TITLE OF INVENTION: PROGNOSIS AND MANAGEMENT OF PROSTATE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/010,398
; FILING DATE: Concurrently Herewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Corder, Timothy S.
; REGISTRATION NUMBER: 38,414
; REFERENCE/DOCKET NUMBER: UROC:015
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512-418-3000


```
; TELEFAX: 512-474-7577
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8224 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-010-398-14

Query Match      4.8%; Score 108.3; DB 1; Length 8224;
Best Local Similarity 59.2%; Pred. No. 1.7e+02;
Matches 218; Conservative 0; Mismatches 147; Indels 3; Gaps 2;

QY 1345 TTTTATCTAGAGTCTTTCTTCTCCAACTATTGTGACAGAAAGTTTCTTAAGTGCA 1404
Db 640 TATTCCTCTGAAAGATTTTATTTCTCTTCATTTCTGAAGGATAGCTTTGCTGAGGGTA 699
QY 1405 GTAGTCTGGCGTACATCTGTAGTCTCTGGAGTCTGTAGCAATCTGTGAGGGCCTTC 1464
Db 700 ATGTTCTTAGCTGACAGTTTCTCTTTTAGTAATTGAATATATTA--TGCCATTTCTTC 757
QY 1465 TTACATTTGAGTTTCTATTGGAAGTCAGGTGTAATCTAATACATCTGCCCTTTATAT 1524
Db 758 CTGGCCTGTAGGTTTCTCCGAGAAATTTCTCTTGTAGTCAGATGGGGGTTACCTTATAT 817
QY 1525 GTTAATTGGCTTTTTCCTTGCACTCTTTTAATATCTTTCTTTGTCTATATCTTTTAG 1584
Db 818 GTGACTTGA-CATTTTCTCTTGCTGCTTTTAAATAATCTTTCTTTGCTTTGACITTTAG 876
QY 1585 TGATTTGATTAATGACATGTGGGGAGTTTCTTTCCGTCCTCAATCTATTGTGGTTTT 1644
Db 877 CAATGCAATTAATAATGCTCTGGAGAGAACCCGTTTAGTGAATTTATTAGGGTTCT 936
QY 1645 GTATGCTTCTGTACCTGTAGGCACTCTTTCTCAAGTTAGGAATTTTCTTTT 1704
Db 937 TATAGCTTCTGTGACCTGGATTTCTATCTCTCTCCAGATGCAAGAAGTTTCTGCTAT 996
QY 1705 GGTITTTCT 1712
Db 997 TATTTTAT 1004

RESULT 40
US-09-366-260-14
; Sequence 14, Application US/09366260
; Patent No. 6171796
; GENERAL INFORMATION:
; APPLICANT: An, Gang
; TITLE OF INVENTION: BIOMARKERS AND TARGETS FOR DIAGNOSIS,
; PROGNOSIS AND MANAGEMENT OF PROSTATE DISEASE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/366,260
; FILING DATE: 03-Aug-1993
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/010,398
; FILING DATE: <unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Corder, Timothy S.
```

```
; REGISTRATION NUMBER: 38,414
; REFERENCE/DOCKET NUMBER: UROC:015
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512-418-3000
; TELEFAX: 512-474-7577
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8224 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-366-260-14

Query Match      4.8%; Score 108.3; DB 1; Length 8224;
Best Local Similarity 59.2%; Pred. No. 1.7e+02;
Matches 218; Conservative 0; Mismatches 147; Indels 3; Gaps 2;

QY 1345 TTTTATCTAGAGTCTTTCTTCTCCAACTATTGTGACAGAAAGTTTCTTAAGTGCA 1404
Db 640 TATTCCTCTGAAAGATTTTATTTCTCTTCATTTCTGAAGGATAGCTTTGCTGAGGGTA 699
QY 1405 GTAGTCTGGCGTACATCTGTAGTCTCTGGAGTCTGTAGCAATCTGTGAGGGCCTTC 1464
Db 700 ATGTTCTTAGCTGACAGTTTCTCTTTTAGTAATTGAATATATTA--TGCCATTTCTTC 757
QY 1465 TTACATTTGAGTTTCTATTGGAAGTCAGGTGTAATCTAATACATCTGCCCTTTATAT 1524
Db 758 CTGGCCTGTAGGTTTCTCCGAGAAATTTCTCTTGTAGTCAGATGGGGGTTACCTTATAT 817
QY 1525 GTTAATTGGCTTTTTCCTTGCACTCTTTTAATATCTTTCTTTGTCTATATCTTTTAG 1584
Db 818 GTGACTTGA-CATTTTCTCTTGCTGCTTTTAAATAATCTTTCTTTGCTTTGACITTTAG 876
QY 1585 TGATTTGATTAATGACATGTGGGGAGTTTCTTTCCGTCCTCAATCTATTGTGGTTTT 1644
Db 877 CAATGCAATTAATAATGCTCTGGAGAGAACCCGTTTAGTGAATTTATTAGGGTTCT 936
QY 1645 GTATGCTTCTGTACCTGTAGGCACTCTTTCTCAAGTTAGGAATTTTCTTTT 1704
Db 937 TATAGCTTCTGTGACCTGGATTTCTATCTCTCTCCAGATGCAAGAAGTTTCTGCTAT 996
QY 1705 GGTITTTCT 1712
Db 997 TATTTTAT 1004

RESULT 41
US-09-702-705-604
; Sequence 604, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 604
; LENGTH: 468
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
```

```

; NAME/KEY: misc.feature
; LOCATION: (1)...(468)
; OTHER INFORMATION: n = A,T,C or G
US-09-702-105-604

Query Match      4.7%; Score 106.2; DB 1; Length 468;
Best Local Similarity 57.7%; Pred. No. 2.9e+03;
Matches 239; Conservative 0; Mismatches 171; Indels 4; Gaps 3;

QY 381 TTGTTGTTGTTTATCTAGATTAAAGCTGGTGGTGCAGATAGACATAGAGTATTATT 440
Db 17 TTCTTAATCCCTGAGTTCTGNTTGAATGCACTGGTCTGAGAGATAGTTTGTATAATT 76

QY 441 TCAATTGCTTTTATCTGTCGAGACTTCTTTGTTTGAATAATGTAATCAATTTTGGAG 500
Db 77 TCTGTTCTTTTACACTTACTGAGGAGAGCTTTACTTCCAGTAATGTCGATTTGGAA 136

QY 501 A--GTTTCATAGGGTCTGCAAGAAGGTACAGTCT--TTGTGTTTGGTGAATAGTCTG 557
Db 137 TAGGTGTGGTGTGCGTCTGAAAGAATGATATATCTGTTGATTTGGGGTGGAGAGTTCTG 196

QY 558 TAAATATCTCTAGGTCCTAGTTCCTGTTTATGACATGATTTAGTCCAGCAATTTCTGTTTC 617
Db 197 TANATGCTCTATTAGTCCGCTTGGTGCAGAGTTGAGTTCAATTCCTGATAGCCTTGTTA 256

QY 618 GTTTTTTGTGAGATGACCTAACTGTTGGAGAGATGGGTATTGAAGTAGCCCATATC 677
Db 257 ACTTTCTGTCGTTGATCTGCTAAATGTTGACAGTGGGTGGAAGTCTCCCATATT 316

QY 678 TGTGTGTGAG-GTCAATATGTCATTTTATGCTGTAGCTGTGCTTGTGTTATGAACCTGGGT 736
Db 317 ATTGTGTGGAGTCTTAAGTCTCTTTGTAGGTCACTAAGGACTTGTCTTTATGAATCTGGGT 376

QY 737 GACATTGTTGTTGGTGCATAGACATTAAGAATTGCAATGTCCTCTTGGTGGATT 790
Db 377 GTCCTGCAATGGGTGCACATATATTTAGGACAGNAGCTCTTCTTGTGAATT 430

RESULT 43
US-09-614-124B-604
; Sequence 604, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Bang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 604
; LENGTH: 468
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(468)
; OTHER INFORMATION: n = A,T,C or G
US-09-614-124B-604

Query Match      4.7%; Score 106.2; DB 1; Length 468;
Best Local Similarity 57.7%; Pred. No. 2.9e+03;
Matches 239; Conservative 0; Mismatches 171; Indels 4; Gaps 3;

QY 381 TTGTTGTTGTTTATCTAGATTAAAGCTGGTGGTGCAGATAGACATAGAGTATTATT 440
Db 17 TTCTTAATCCCTGAGTTCTGNTTGAATGCACTGGTCTGAGAGATAGTTTGTATAATT 76

QY 441 TCAATTGCTTTTATCTGTCGAGACTTCTTTGTTTGAATAATGTAATCAATTTTGGAG 500
Db 77 TCTGTTCTTTTACACTTACTGAGGAGAGCTTTACTTCCAGTAATGTCGATTTGGAA 136

QY 501 A--GTTTCATAGGGTCTGCAAGAAGGTACAGTCT--TTGTGTTTGGTGAATAGTCTG 557
Db 137 TAGGTGTGGTGTGCGTCTGAAAGAATGATATATCTGTTGATTTGGGGTGGAGAGTTCTG 196

QY 558 TAAATATCTCTAGGTCCTAGTTCCTGTTTATGACATGATTTAGTCCAGCAATTTCTGTTTC 617
Db 197 TANATGCTCTATTAGTCCGCTTGGTGCAGAGTTGAGTTCAATTCCTGATAGCCTTGTTA 256

QY 618 GTTTTTTGTGAGATGACCTAACTGTTGGAGAGATGGGTATTGAAGTAGCCCATATC 677
Db 257 ACTTTCTGTCGTTGATCTGCTAAATGTTGACAGTGGGTGGAAGTCTCCCATATT 316

QY 678 TGTGTGTGAG-GTCAATATGTCATTTTATGCTGTAGCTGTGCTTGTGTTATGAACCTGGGT 736
Db 317 ATTGTGTGGAGTCTTAGTCTCTTTGAGTCACTAAGGACTTGTCTTTATGAATCTGGGT 376

QY 737 GACATTGTTGTTGGTGCATAGACATTAAGAATTGCAATGTCCTCTTGGTGGATT 790
Db 377 GTCCTGCAATGGGTGCACATATATTTAGGACAGNAGCTCTTCTTGTGAATT 430

RESULT 42
US-09-736-457-604
; Sequence 604, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 604
; LENGTH: 468
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(468)
; OTHER INFORMATION: n = A,T,C or G
US-09-736-457-604

Query Match      4.7%; Score 106.2; DB 1; Length 468;
Best Local Similarity 57.7%; Pred. No. 2.9e+03;
Matches 239; Conservative 0; Mismatches 171; Indels 4; Gaps 3;

```

```
Db 137 TAGGTGCGTCTCGTCTGAAAGAAATGTAATTTCTGTGATTTGGGGTGAGAGTTCTG 196
Qy 558 TAAATATCTCTAGGTCCACTTGGTTTATGACATCAGTTAGCTCCAGCATTTCTCTGTTTC 617
Db 197 TANATGCTATTAGGTCGCTGGTGCAGAGTTGAGTTCAATTCCTGGATAGCCTTGTTA 256
Qy 618 GTTTTCTGTTGAGATGACCTAACTGTTGGAGAGAAATGGGGTATTGAAGTAGCCCACTATC 677
Db 257 ACTTTCTGCTCGTTGATCTGCTAAATGTTGACAGTGGGGTGAAGAGTCTCCCAATATT 316
Qy 678 TGTGTTGTCAG-GTCAATATGATTTAGCTAGCTGCTGCTGTTTATGAACTTCGGT 736
Db 317 ATTGTTGGGAGTCTAGTCTCTTTGTAGTCACTAAGGACTTCTCTTATGAATCTGGGT 376
Qy 737 GACATTGTTTGGTGCATAGACATTAAGAATTGCAATGCTCTCTGGTGGATT 790
Db 377 GCTCCTGCATTGGGTGCACATATATTAGGACAGCNAGCTCTTCTTGTGAATT 430

RESULT 44
US-09-671-325-604
; Sequence 604, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478612
; CURRENT APPLICATION NUMBER: US/09/671.325
; CURRENT FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 1325
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 604
; LENGTH: 468
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(468)
; OTHER INFORMATION: n = A,T,C or G
US-09-671-325-604

Query Match 4.7%; Score 106.2; DB 1; Length 468;
Best Local Similarity 57.7%; Pred. No. 2.9e+03;
Matches 239; Conservative 0; Mismatches 171; Indels 4; Gaps 3;

Qy 381 TTGTTGTTGTTTATCTAGATTTAAAGCTGTGGTGCAGATAGACATAGATTAATT 440
Db 17 TTCTTAATCCTGAGTCTCGNTTGATTGCACCTGTGGTCTGAGAGATAGTTTGTATAATT 76
Qy 441 TCAATTTGCTTTTATCTGTCGAGACTTGTCTTTGTTTGAATATGTAATTTTGGAG 500
Db 77 TCTGTTCTTTTACACTTACTGAGGAGAGCTTTACTTCCAAGTATGTTGTCGATTTGGAA 136
Qy 501 A--GTTTCAATAGGCTGCTGACAAAGGTAACAGTCT-TTGTGTTTGGTGAATAGTCTG 557
Db 137 TAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 196
Qy 558 TAAATATCTCTAGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 617
Db 197 TANATGCTATTAGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 256
Qy 618 GTTTTCTGTTGATGACCTAACTGTTGGAGAGAAATGGGGTATTGAAGTAGCCCACTATC 677
Db 257 ACTTTCTGCTCGTTGATCTGCTAAATGTTGACAGTGGGGTGAAGAGTCTCCCAATATT 316
Qy 678 TGTGTTGTCAG-GTCAATATGATTTAGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 736
Db 317 ATTGTTGGGAGTCTAAGTCTCTTTGTAGGTCACAAAGGACTTGTCTTATGAATCTGGGT 376
Qy 737 GACATTGCTTTGGTGCATAGACATTAAGAATTGCAATGCTCTCTTGTGGGATT 790
Db 377 GCTCCTGCATTGGGTGCACATATATTAGGACAGCNAGCTCTTCTTGTGAATT 430
```

```
Db 257 ACTTTCTGCTCGTTGATCTGCTAATGTTGACAGTGGGGTGAAGTCTCCCAATATT 316
Qy 678 TGTGTTGTCAG-GTCAATATGATTTAGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 736
Db 317 ATTGTTGGGAGTCTAAGTCTCTTTGTAGGTCACAAAGGACTTGTCTTATGAATCTGGGT 376
Qy 737 GACATTGCTTTGGTGCATAGACATTAAGAATTGCAATGCTCTCTTGTGGGATT 790
Db 377 GCTCCTGCATTGGGTGCACATATATTAGGACAGCNAGCTCTTCTTGTGAATT 430

RESULT 45
US-09-589-184-604
; Sequence 604, Application US/09589184
; Patent No. 6686447
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478612
; CURRENT APPLICATION NUMBER: US/09/589.184
; CURRENT FILING DATE: 2000-06-05
; NUMBER OF SEQ ID NOS: 827
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 604
; LENGTH: 468
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(468)
; OTHER INFORMATION: n = A,T,C or G
US-09-589-184-604

Query Match 4.7%; Score 106.2; DB 1; Length 468;
Best Local Similarity 57.7%; Pred. No. 2.9e+03;
Matches 239; Conservative 0; Mismatches 171; Indels 4; Gaps 3;

Qy 381 TTGTTGTTGTTTATCTAGATTTAAAGCTGTGGTGCAGATAGACATAGATTAATT 440
Db 17 TTCTTAATCCTGAGTCTCGNTTGATTGCACCTGTGGTCTGAGAGATAGTTTGTATAATT 76
Qy 441 TCAATTTGCTTTTATCTGTCGAGACTTGTCTTTGTTTGAATATGTAATTTTGGAG 500
Db 77 TCTGTTCTTTTACACTTACTGAGGAGAGCTTTACTTCCAAGTATGTTGTCGATTTGGAA 136
Qy 501 A--GTTTCAATAGGCTGCTGACAAAGGTAACAGTCT-TTGTGTTTGGTGAATAGTCTG 557
Db 137 TAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 196
Qy 558 TAAATATCTCTAGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 617
Db 197 TANATGCTATTAGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 256
Qy 618 GTTTTCTGTTGATGACCTAACTGTTGGAGAGAAATGGGGTATTGAAGTAGCCCACTATC 677
Db 257 ACTTTCTGCTCGTTGATCTGCTAAATGTTGACAGTGGGGTGAAGAGTCTCCCAATATT 316
Qy 678 TGTGTTGTCAG-GTCAATATGATTTAGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 736
Db 317 ATTGTTGGGAGTCTAAGTCTCTTTGTAGGTCACAAAGGACTTGTCTTATGAATCTGGGT 376
Qy 737 GACATTGCTTTGGTGCATAGACATTAAGAATTGCAATGCTCTCTTGTGGGATT 790
Db 377 GCTCCTGCATTGGGTGCACATATATTAGGACAGCNAGCTCTTCTTGTGAATT 430
```

us10664775-4.rni

Wed Aug 25 15:11:59 2004

Search completed: August 25, 2004, 08:12:44
Job time : 3688 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2004, 08:14:16 ; Search time 4409 Seconds
(without alignments)
3.936 Million cell updates/sec

Title: US-10-664-775-4

Perfect score: 2279

Sequence: 1 gatcaactctcttagtgaag.....ttgtaattcttagtgctgat 2279

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 0.5

Searched: 41 seqs, 3807569 residues

Total number of hits satisfying chosen parameters: 82

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : rnpb4.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2279	100.0	2279	1	US-10-664-775-4
2	2279	100.0	2715	1	US-10-664-775-1
3	2251	98.8	2267	1	US-10-664-775-5
4	2003	87.9	2003	1	US-10-664-775-3
5	646.9	27.5	41219	1	US-10-087-192-577
6	610.3	26.8	193853	1	US-10-087-192-1663
7	591.499	26.0	13619	1	US-10-369-214-129
8	590.199	25.9	178825	1	US-10-087-192-1945
9	588.5	25.8	6335	1	US-10-388-934-580
10	588.5	25.8	6335	1	US-10-388-934-685
11	588.5	25.8	6335	1	US-10-388-934-790
12	588.3	25.8	6335	1	US-10-191-803-720
13	587.8	25.8	29604	1	US-10-374-077-207
14	587.8	25.8	283190	1	US-10-322-281-115
15	585.399	25.7	225883	1	US-10-175-523-57
16	585.1	25.7	173808	1	US-10-003-806-10
17	580.301	25.5	120239	1	US-10-322-281-271
18	578.299	25.4	173808	1	US-10-003-806-10
19	578.2	25.4	8048	1	US-10-388-934-223
20	578.2	25.4	8048	1	US-10-388-934-224
21	578.2	25.4	8048	1	US-10-388-934-682
22	578.2	25.4	8048	1	US-10-388-934-720
23	578.2	25.4	8048	1	US-10-388-934-785
24	578.2	25.4	8048	1	US-10-388-934-808
25	574.999	25.2	116858	1	US-10-087-192-547
26	573.4	25.2	200400	1	US-10-087-192-1033
27	572.9	25.1	225883	1	US-10-175-523-57
28	570.9	25.1	95982	1	US-10-087-192-2029
29	567.5	24.9	251364	1	US-10-175-523-58
30	567.5	24.9	251364	1	US-10-175-523-61
31	567.5	24.9	251364	1	US-10-175-523-79
32	566.4	24.9	175338	1	US-10-087-192-1861
33	565.399	24.8	51259	1	US-10-374-077-209

ALIGNMENTS

RESULT 1

```
US-10-664-775-4      1  US-10-087-192-1861      Sequence 1861, Ap
35 558.6      24.5 176594 1  US-10-322-281-495      Sequence 495, Ap
36 556.899    24.4 85548  1  US-10-175-523-75      Sequence 75, Ap
37 555.699    24.4 47115  1  US-10-052-482-133     Sequence 133, Ap
38 550.599    24.2 68370  1  US-10-087-192-1765     Sequence 1765, Ap
39 547.299    24.0 106664 1  US-10-175-523-97      Sequence 97, Ap
40 546.899    24.0 104399 1  US-10-322-281-123     Sequence 123, Ap
c 41 546.799    24.0 185548 1  US-10-175-523-62      Sequence 62, Ap
42 537.898    23.6 91071  1  US-10-087-192-235     Sequence 235, Ap
43 525.898    23.1 96597  1  US-10-052-482-7       Sequence 7, Ap
44 521.797    22.9 51259  1  US-10-374-077-209     Sequence 209, Ap
c 45 521.599    22.9 96602  1  US-10-085-117-61      Sequence 61, Ap

GENERAL INFORMATION:
; APPLICANT: Simesen, Ruth B
; APPLICANT: Pedersen, Anette A
; APPLICANT: Paisst, Steffen
; APPLICANT: Jensen, Jan J
; APPLICANT: Weilguny, Dietmar
; TITLE OF INVENTION: Method for Making Recombinant Proteins
; FILE REFERENCE: 6448.200-US
; CURRENT APPLICATION NUMBER: US/10/664,775
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: Danish Application PA 2002 01384
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US 60/416,566
; PRIOR FILING DATE: 2002-10-07
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2279
; TYPE: DNA
; ORGANISM: Cricetulus griseus
US-10-664-775-4

Query Match      100.0%; Score 2279; DB 1; Length 2279;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 2279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATCACTCTCTAGTCAAAAGGTGGGGGTCTGAGGCTCCAAATGCTGTTGATGTTGTTGAG 60
Db 1 GATCACTCTCTAGTCAAAAGGTGGGGGTCTGAGGCTCCAAATGCTGTTGATGTTGTTGAG 60
Qy 61 TATCTCATACAGAGGATAGCAGTATGCTGCTGGACATAGGTAAGCTTTCCAGAGAG 120
Db 61 TATCTCATACAGAGGATAGCAGTATGCTGCTGGACATAGGTAAGCTTTCCAGAGAG 120
Qy 121 ACTTCATAATATATTTTCTTGAAGCTCTGCTGGCAATACCTCTGGGGGTCTGCTTTTC 180
Db 121 ACTTCATAATATATTTTCTTGAAGCTCTGCTGGCAATACCTCTGGGGGTCTGCTTTTC 180
Qy 181 TCCTGCTGATTCCTAGGCTGAGGTTACCACTGCTCTCTCTCCCTTTCTCTAACAC 240
Db 181 TCCTGCTGATTCCTAGGCTGAGGTTACCACTGCTCTCTCTCTCTCTCTCTCTCTAACAC 240
Qy 241 TTCTGGCCAGGGTAGGGGCACTACCGCAATTCCTCTCTCTTCCAAACACTTCTATTTC 300
Db 241 TTCTGGCCAGGGTAGGGGCACTACCGCAATTCCTCTCTCTCTCTCTCTCTCTATTTC 300
Qy 301 TGAATTTCTATCTGGCTCAATTTTAACTCAGTAGTGAGTTGTTGTTTCCATAAGTTTG 360
Db 301 TGAATTTCTATCTGGCTCAATTTTAACTCAGTAGTGAGTTGTTGTTTCCATAAGTTTG 360
Qy 361 TGAATTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 420
```



```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 2715
; TYPE: DNA
; ORGANISM: Cricetulus griseus
US-10-664-775-1

Query Match      100.0%; Score 2279; DB 1; Length 2715;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 2279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  GATCAGCTCTCTAGTGAAGGTGGGGGTCTCAGAGCTCCAAATGGTGTGTGATGGTAGAG 60
DB      437  GATCAGCTCTCTAGTGAAGGTGGGGGTCTCAGAGCTCCAAATGGTGTGTGATGGTAGAG 496
QY      61  TATCTCATACAGAGATAGCACTAGATGCTGTCTGGGACATAGTAAGCTTTCCAGAGAG 120
DB      497  TATCTCATACAGAGATAGCACTAGATGCTGTCTGGGACATAGTAAGCTTTCCAGAGAG 556
QY      121  ACTTCATAATATATTTCTTCAAGCCCTGCTGGCAATACCTCTGGGGCTGCTGCCCTTC 180
DB      557  ACTTCATAATATATTTCTTCAAGCCCTGCTGGCAATACCTCTGGGGCTGCTGCCCTTC 616
QY      181  TCCCTGTCTGATTCCTAGGGTGGAGGTACCAGCTGCTCTCTCTCTCCCTTTCTCTAACAC 240
DB      617  TCCCTGTCTGATTCCTAGGGTGGAGGTACCAGCTGCTCTCTCTCTCCCTTTCTCTAACAC 676
QY      241  TTCTGGGCCAGGTAGGGGCACTACCGCAATTCCTCTCTCTCTCTCCAAACACTTCTATTCT 300
DB      677  TTCTGGGCCAGGTAGGGGCACTACCGCAATTCCTCTCTCTCTCTCCAAACACTTCTATTCT 736
QY      301  TGATTTCTATCTTGCTCATTTTAACTCAGTAGTAGTGTGTTGGTTTCCATAAGTTTG 360
DB      737  TGATTTCTATCTTGCTCATTTTAACTCAGTAGTAGTGTGTTGGTTTCCATAAGTTTG 796
QY      361  TAAGTTTCTGTGTTCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 420
DB      797  TAAGTTTCTGTGTTCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 856
QY      421  ATAGACATAGAGATATTTCAATGTCCTTTTATCTGTCGAGACTTGTCTTGTGTTGAA 480
DB      857  ATAGACATAGAGATATTTCAATGTCCTTTTATCTGTCGAGACTTGTCTTGTGTTGAA 916
QY      481  ATATGTATTCATTTTGAGAGATTTTCATAGGGTGTCTGACAGAGAGGTACAGTCTTTGT 540
DB      917  ATATGTATTCATTTTGAGAGATTTTCATAGGGTGTCTGACAGAGAGGTACAGTCTTTGT 976
QY      541  TTGSGTGAATAGTCTGAATATCTCTAGGTCACCTTGGTTTATGACATCAGTAGCTC 600
DB      977  TTGSGTGAATAGTCTGAATATCTCTAGGTCACCTTGGTTTATGACATCAGTAGCTC 1036
QY      601  CAGCATTTCTCTGTTTCGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 660
DB      1037  CAGCATTTCTCTGTTTCGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1096
QY      661  TGAAGTAGCCCACTATCTGTGTGTGAGGTCAATATGATGATTTTATGCTGTAGCTGTG 720
DB      1097  TGAAGTAGCCCACTATCTGTGTGTGAGGTCAATATGATGATTTTATGCTGTAGCTGTG 1156
QY      721  TTTTATGACCTGGGTGACATTTGTTGGTGGTGCATAGACATTAAGATTGAATGTCCTC 780
DB      1157  TTTTATGACCTGGGTGACATTTGTTGGTGGTGCATAGACATTAAGATTGAATGTCCTC 1216
QY      781  TTGTTGGATTTTCTTTGATGCTATGATGATTTCTCCCAATCTCATCTGCTTAGTTT 840
DB      1217  TTGTTGGATTTTCTTTGATGCTATGATGATTTCTCCCAATCTCATCTGCTTAGTTT 1276
QY      841  GGGTTTATGCTATTTAGTCAGATATTAATAAGACTGTATCGGCTGCTTTCTTAGGGCAT 900
DB      1277  GGGTTTATGCTATTTAGTCAGATATTAATAAGACTGTATCGGCTGCTTTCTTAGGGCAT 1336
QY      901  TTGCTTAGAATATCTTTTCCATCTTTTACCTTAAGGTGATGCTATCCATGGTAGGTTG 960
```

QY 2041 AGTATTTACATTTTCATAGGTTTCTTTAATGAATTAATTAATTCCTCTTCAAGGACCTT 2100
Db 2477 AGTATTTACATTTTCATAGGTTTCTTTAATGAATTAATTAATTCCTCTTCAAGGACCTT 2536
QY 2101 TTATGAATTCATAAATATGATGTTAAGTCTCTTGGCTTGCTTCAGCTATGTTGCAATTC 2160
Db 2537 TTATGAATTCATAAATATGATGTTAAGTCTCTTGGCTTGCTTCAGCTATGTTGCAATTC 2596
QY 2161 TCAGGGCCCTATTGTAATAGGTTTTCAGCAGGACATATTGTCCTGGTGTGTTATTGTCGT 2220
Db 2597 TCAGGGCCCTATTGTAATAGGTTTTCAGCAGGACATATTGTCCTGGTGTGTTATTGTCGT 2656
QY 2221 GTTTTGTCTTTGGCATATAGACGGCTGAGTTTGGGATGATGTAATTCATAGGTCCTGAT 2279
Db 2657 GTTTTGTCTTTGGCATATAGACGGCTGAGTTTGGGATGATGTAATTCATAGGTCCTGAT 2715

RESULT 3
US-10-664-775-5
; Sequence 5, Application US/10664775
; Publication No. US20040115776A1
; GENERAL INFORMATION:
; APPLICANT: Simesen, Ruth B
; APPLICANT: Pedersen, Anette A
; APPLICANT: Falsst, Steffen
; APPLICANT: Jensen, Jan J
; APPLICANT: Weiliguny, Dietmar
; TITLE OF INVENTION: Method for Making Recombinant Proteins
; FILE REFERENCE: 6448.200-US
; CURRENT APPLICATION NUMBER: US/10/664,775
; CURRENT FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: Danish Application PA 2002 01384
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US 60/416,566
; PRIOR FILING DATE: 2002-10-07
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 5
; LENGTH: 2267
; TYPE: DNA
; ORGANISM: Cricetulus griseus
US-10-664-775-5

Query Match 98.8%; Score 2251; DB 1; Length 2267;
Best Local Similarity 99.5%; Pred. No. 1.4e-06;
Matches 2267; Conservative 0; Mismatches 0; Indels 12; Gaps 1;

QY 1 GATCACTCCTCTAGTGAAGGTGGGGTCTGAGGCTCCAATGGTGTGATGGGTAGAG 60
Db 1 GATCACTCCTCTAGTGAAGGTGGGGTCTGAGGCTCCAATGGTGTGATGGGTAGAG 60
QY 61 TATCTCATACAGAGATAGCACTAGATGCTCTCTGGGACATAGTAAGCTTCCAGAGAG 120
Db 61 TATCTCATACAGAGATAGCACTAGATGCTCTCTGGGACATAGTAAGCTTCCAGAGAG 120
QY 121 ACTTCATAATATATTTCTTGAAGCCTCTGCTGGCAATACCTTCTGGGGCTGCTGCCCTTC 180
Db 121 ACTTCATAATATATTTCTTGAAGCCTCTGCTGGCAATACCTTCTGGGGCTGCTGCCCTTC 180
QY 181 TCCCTGCTGATTCCTAGGTTGAGGGTTACCACTGCTCTCTCTCCCTTCTCTTAACAC 240
Db 181 TCCCTGCTGATTCCTAGGTTGAGGGTTACCACTGCTCTCTCTCCCTTCTCTTAACAC 240
QY 241 TTCTGGGGCAGGCTAGGGGCACTACCGCATTCCTCTCTCTTCCAAACACTTCTATTTC 300
Db 241 TTCTGGGGCAGGCTAGGGGCACTACCGCATTCCTCTCTCTTCCAAACACTTCTATTTC 300
QY 301 TGATTTCTATCTGGCTCATTTTAACTCAGTAGTAGGTTGTTGGTTTCCATAAGTTTG 360
Db 301 TGATTTCTATCTGGCTCATTTTAACTCAGTAGTAGGTTGTTGGTTTCCATAAGTTTG 360
QY 361 TAAGTTTCTGTTCTGTTGTTGTTGTTGTTATCTAGATTAAAGCTGTGGTGGTCAG 420

Db 361 TAAGTTTCTGTTCTGTTGTTGTTGTTGTTATCTAGATTAAAGCTGTGGTGGTCAG 420
QY 421 ATAGGACATAGAGTATTAATTTCAATTTGCTTTATCTGCGAGACTTCTGTTGTTTCAA 480
Db 421 ATAGGACATAGAGTATTAATTTCAATTTGCTTTATCTGCGAGACTTCTGTTGTTGAA 480
QY 481 ATATGTAATTTCAATTTTGGAGAGTTTTCATAGGGTGTGCAAGAAAGGTACAGTCTTTG 540
Db 481 ATATGTAATTTCAATTTTGGAGAGTTTTCATAGGGTGTGCAAGAAAGGTACAGTCTTTG 540
QY 541 TTTGGTGAATAGTCTGTAATATCTCTAGGTCACCTTGGTTTATGACATCAGTTAGCTC 600
Db 541 TTTGGTGAATAGTCTGTAATATCTCTAGGTCACCTTGGTTTATGACATCAGTTAGCTC 600
QY 601 CAGCAATTTCTGTTTTCGTTTTCGTTTTCGTTTTCGTTTTCGTTTTCGTTTTCGTTT 660
Db 601 CAGCAATTTCTGTTTTCGTTTTCGTTTTCGTTTTCGTTTTCGTTTTCGTTTTCGTTT 660
QY 661 TGAAGTAGCCCACTATCTGTTGTTGAGGTCAATATGATGATTTAGCTGTAGCTGTCTTG 720
Db 661 TGAAGTAGCCCACTATCTGTTGTTGAGGTCAATATGATGATTTAGCTGTAGCTGTCTTG 720
QY 721 TTTTATGAACCTTTGGGTGACATTTGTTGGTGCATAGACATTAAGAAATGCAATGCTC 780
Db 721 TTTTATGAACCTTTGGGTGACATTTGTTGGTGCATAGACATTAAGAAATGCAATGCTC 780
QY 781 TTGGTGAATTTTCCCTTTGATGCTATGATGATTTTCCCAATCTCATCTGCTTGTAGTT 840
Db 781 TTGGTGAATTTTCCCTTTGATGCTATGATGATTTTCCCAATCTCATCTGCTTGTAGTT 840
QY 841 GGGTTTAACTTATTTAGTCAGATATTAAGTGCATGATCGGCTTCTCTTAGGCGCAT 900
Db 841 GGGTTTAACTTATTTAGTCAGATATTAAGTGCATGATCGGCTTCTCTTAGGCGCAT 900
QY 901 TTGCTTAGAATATCTTTTCCATCTTTTAACTTAACTTAACTTAACTTAACTTAACTTAA 960
Db 901 TTGCTTAGAATATCTTTTCCATCTTTTAACTTAACTTAACTTAACTTAACTTAACTTAA 960
QY 961 TCTTTTTCGATCGACGACGATAGGATGATGATGATGATGATGATGATGATGATGATGAT 1020
Db 961 TCTTTTTCGATCGACGACGATAGGATGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 1021 TCTTTTTCGATCGACGACGATAGGATGATGATGATGATGATGATGATGATGATGATGAT 1080
Db 1021 TCTTTTTCGATCGACGACGATAGGATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1081 TGTGTTGAGATTTCTGTTTATCTTTCGACTTTCGAAAGTGTGTTGTTGTTGTTGTTGTT 1140
Db 1081 TGTGTTGAGATTTCTGTTTATCTTTCGACTTTCGAAAGTGTGTTGTTGTTGTTGTTGTT 1128
QY 1141 GTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1200
Db 1129 GTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1188
QY 1201 TGTGTTCTCTCCCTCTTTTGAATTTTGGCTTGGAAATTAATTTATTTATTTATTTTCTTG 1260
Db 1189 TGTGTTCTCTCCCTCTTTTGAATTTTGGCTTGGAAATTAATTTATTTATTTATTTTCTTG 1248
QY 1261 AATGTTGGGTAAACATCTTATAGATTGAAGTTTTCCTAGCTTCTTTTAGGTTCTGATTTG 1320
Db 1249 AATGTTGGGTAAACATCTTATAGATTGAAGTTTTCCTAGCTTCTTTTAGGTTCTGATTTG 1308
QY 1321 AAGATAGATATTTCTTACATCTGATTTTATCTTAGAATTTCTTTCTTTCTCCAACTATTG 1380
Db 1309 AAGATAGATATTTCTTACATCTGATTTTATCTTAGAATTTCTTTCTTTCTCCAACTATTG 1368
QY 1381 TCACAGAAAGTTTTCAGTACGATGCTGGCTGACATCTGATGCTCTTGGAGTCT 1440
Db 1369 TCACAGAAAGTTTTCAGTACGATGCTGGCTGACATCTGATGCTCTTGGAGTCT 1428
QY 1441 GTAGCACATCTGTGAGGGCCCTTCTTACATTTTGAATTTTCTATTGGAAAAGTCAGGTGA 1500

Db 901 TCTGTTACCCAGTATCTTTTCTAGAGAAAATAAGATCAATTGATGATGATGAGAA 960
Qy 1067 TTATCAATGAGCAGTGTGGGATCTTGTTATCTTGCACTGTGAGTGTGTGTGT 1126
Db 961 TTATCAATGAGCAGTGTGGGATCTTGTTATCTTGCACTGTGAGTGTGTGTGT 1020
Qy 1127 GT 1186
Db 1021 GT 1080
Qy 1187 GT 1246
Db 1081 GT 1140
Qy 1247 TTCAATATTTCTTGAATGTGGTAACATCTTTAGATTGAAGTGTGTGTGTGTGT 1306
Db 1141 TTCAATATTTCTTGAATGTGGTAACATCTTTAGATTGAAGTGTGTGTGTGTGT 1200
Qy 1307 TAGGTCTGCATTTGAAGATAGATATCTTTTACATCTGATTTTATCTTAGAATGCTTTCT 1366
Db 1201 TAGGTCTGCATTTGAAGATAGATATCTTTTACATCTGATTTTATCTTAGAATGCTTTCT 1260
Qy 1367 TTCTCCAACTATGTGACAGAAAGTTTTCTAAGTGCACTAGTCTGGCTGACATCTGTA 1426
Db 1261 TTCTCCAACTATGTGACAGAAAGTTTTCTAAGTGCACTAGTCTGGCTGACATCTGTA 1320
Qy 1427 GTCTCTTGAGTCTGTAGCACATCTGTGACGGGCTTCTTACATTTTGAATTTTCTATTGG 1486
Db 1321 GTCTCTTGAGTCTGTAGCACATCTGTGACGGGCTTCTTACATTTTGAATTTTCTATTGG 1380
Qy 1487 AAAAGTCAAGTGTAAATCTTAATACATCTGCGCTTATATGTTAAATGGTCTTTTCCCTT 1546
Db 1381 AAAAGTCAAGTGTAAATCTTAATACATCTGCGCTTATATGTTAAATGGTCTTTTCCCTT 1440
Qy 1547 GCATCTTTTAAATATCTTTCTTTGTTCTATCTTTTGTAGTATTTGATTTATATGCACTGT 1606
Db 1441 GCATCTTTTAAATATCTTTCTTTGTTCTATCTTTTGTAGTATTTGATTTATATGCACTGT 1500
Qy 1607 GGGAGTCTTTTCCGCTCCAACTATTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1666
Db 1501 GGGAGTCTTTTCCGCTCCAACTATTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1560
Qy 1667 GGCATCTCTTTCTCAAGGTAGGAAATTTTCTTTTGGTGTGTGTGTGTGTGTGTGTGTGT 1726
Db 1561 GGCATCTCTTTCTCAAGGTAGGAAATTTTCTTTTGGTGTGTGTGTGTGTGTGTGTGTGT 1620
Qy 1727 CTGCTTTTGACGTGCTTTTCCCTTCCCTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1786
Db 1621 CTGCTTTTGACGTGCTTTTCCCTTCCCTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1680
Qy 1787 GGCCTTCCGGATGTTTATGCTCGATTTATTTAGACTTAACATTTTCTTTGACCAAGGT 1846
Db 1681 GGCCTTCCGGATGTTTATGCTCGATTTATTTAGACTTAACATTTTCTTTGACCAAGGT 1740
Qy 1847 ATCCATTTCTCTATCTTGTCTTCACTGCTGAGATTTCTCTTCTTCTTCTTCTTCTTCTTCT 1906
Db 1741 ATCCATTTCTCTATCTTGTCTTCACTGCTGAGATTTCTCTTCTTCTTCTTCTTCTTCTTCT 1800
Qy 1907 GTCACTGAGGCTGTCTGAGGTTCCGTTGGTGTCTTAAATTTTCTTCTTCTTCTTCTTCTTCT 1966
Db 1801 GTCACTGAGGCTGTCTGAGGTTCCGTTGGTGTCTTAAATTTTCTTCTTCTTCTTCTTCTTCT 1860
Qy 1967 CTTTCAGTTTGGGTTTGTGTTTAAATTTCTATTTTCCACTTTCAGGTCTCTGAAATGTTTAA 2026
Db 1861 CTTTCAGTTTGGGTTTGTGTTTAAATTTCTATTTTCCACTTTCAGGTCTCTGAAATGTTTAA 1920
Qy 2027 CTCATTTTCCCTCCAGTATTTTACATTTTCAATGAGTGTCTTTTAAATGAGTATTTATTCATTC 2086
Db 1921 CTCATTTTCCCTCCAGTATTTTACATTTTCAATGAGTGTCTTTTAAATGAGTATTTATTCATTC 1980
Qy 2087 TCTTCAAGGACCTTTTATGAATT 2109

Db 1981 TCTTCAAGGACCTTTTATGAATT 2003
RESULT 5
US-10-087-192-577
; Sequence 577, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 577
; LENGTH: 41219
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(41219)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-577
Query Match 27.5%; Score 626.9; DB 1; Length 41219;
Best Local Similarity 65.8%; Pred. No. 1.8; Gaps 23;
Matches 1223; Conservative 0; Mismatches 476; Indels 161;
Qy 292 TCTATTTCTTGATTTTCTATCTTTGGCTCAATTTTAACTCAGTAGTGTGAGTTGTTTCC 351
Db 37261 TTATTTCTTTATTTCTTCCCTGACTCAGAGATCATTTGAGAAGGGTGTTCATTTCC 37320
Qy 352 ATAAGTTCTGTAAGTTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTT 411
Db 37321 ATGAATTTCTGAGCTTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTT 37377
Qy 412 GGTGGTCAGATAGGACATAGAGTATTTTCAATTTGTTCTTTTATCTGTCGAGACTTGCTT 471
Db 37378 GATGGTCTGTAAGATACAGGCAATTTTCAATTTCTTTGTTCTGTTGAGCTTGCTT 37437
Qy 472 TGTTTTGAATATGTAATTCATTTTGGAGA--GTTTCATAGGGTCTGACAAAGAGTA- 528
Db 37438 TGGACCAACTATATGTTGAGTTCAGTTTGGAGAGGATCCATGAGTGTCTGAAAAGAGTAT 37497
Qy 529 --CAGTCTTTGTTTGGTGAATAGTCTGTAATATCT-CTAGTCCACTTGTGTTAT 585
Db 37498 TTTTCTTTTATGTTTGGTGAATATTTCCATGAGTGTCTATAGTCCATTTCAATTCAT 37557
Qy 586 GACATCAGTGTAGTCCAGCAATTTCTGTTTCTGTTTCTGTTTGTGAGATGACCTPAATGTTG 645
Db 37558 AATGTCTACTAGTTTCTTTATTTCTGTTTGTAG-TTTCTGCTCGATGACATGTCAATTG 37616
Qy 646 GAGAGATCGGGTATGGAAGTAGCCCACTATCTGTTGTGAGG-TCAATATGTTGATTTTA 704
Db 37617 GTGAGTAGG----TGTTGAAGTCTCTGTTAAATGTGTTGGGTTCAAGTGTGATTTAA 37671
Qy 705 GCTGTAGTGTGTTGTTGTTTATGAACCTTGGGTGACATTTGTGTTGTGTCATAGACATTAA 764
Db 37672 GCTTTAGCAGTGTTCCTTTTATAGCTGTGGTGTGCTTGTGTTGGGCATAGATTTCA 37731
Qy 765 GAATTCGAATGCTCTTGTGTGA-TTTTCTTTGATGCCATGATAGTATTTCTTCCCAAT 823
Db 37732 GAACAGATATCATCTTGGCAGATTTTCTTATATGAGTGTGAGTGTCTTCTTACTGT 37791
Qy 824 CTCATCTGCTTAGTTTGGTGTAAAGTCTTATTAGTCAGATATTAATAATGACTGTATCGGC 883

Db 37792 CTCCTTTGATTAATTTGGTTGAAAGTTTATCTTTTAGAATTAAGATGGTTACTCCAGC 37851
Qy 884 TTGCTCTTAGGGCCATTTGCTTAGAATA-TCTTTTCCATCCTTTTACTCTTAAGGTGATG 942
Db 37852 TTGCTCTTTGGTCTGTTTGGTTGGAAAACTTTTTCACCCCTTTACTCTCGATAATG 37911
Qy 943 TCTA---TCCATGTAGTGTCTTTTGTGATCAGCAGTAGATGGAICTGTTTTC 999
Db 37912 TCTATTTTATGCTGAGTGTGTTCTCTGACAGAAATTAATGATGGATCCCTGTTTCA 37971
Qy 1000 TATCCATCTCTTACCCAGTATCTTTTCTAGAGAAATTAAGATCAITGAGTCATTGATG 1059
Db 37972 CATTCACCTCTGTTAGCCGTGCTTTTATTTGGGAATGAG-----TCCACTGGTG 38023
Qy 1060 TTGAGAAATTAATGAGCAGTGTGTTGGGATCTCTGTTATCTTGACCTGTGAGGTG 1119
Db 38024 TTGAGAGATTAATAGCAGTGTGTTTATTTTACAGGTATTTTCTGCTGTGAGTGTATC 38083
Qy 1120 TGT 1179
Db 38084 TGT 38117
Qy 1180 TGT 1239
Db 38118 -----CTTGTCTGTCTTTTCTGCCCTCTTTTGAATTTTGGCCTGGAATTA 1239
Qy 1240 TTTTATTTCATATTTCTTGAATGCGGTAAACATCTTTAGATTAAGTATTTTCTCCTAG 1299
Db 38147 CTTAGTACTAGTTTCTTGGATGTAGTATCTCTCTGAGTGTGGAATTTCTCCTCTAG 38206
Qy 1300 CTTCTCTT--TAGGTCTGCAATTTGAAGTAGATATCTTTTACATCTGATTTTATCTTAGAA 1357
Db 38207 AATTTTGGTAGGCTGCAATTTGGCTAGATATGCTTAAATTTGGAT----- 38255
Qy 1358 TGTCTCTTCTTCCACTATTTGTGACAGAAAGTTTCTTAAAGTCAAGTGTGCGCTG 1417
Db 38256 -----CTATGGTAATTTGAGAGCTTTGCTGAGTATAGTGTCTAGTGTG 38298
Qy 1418 ACATCTGAGTCTCTTGA-TGCTGTAGCAGCATCTGTGCGAGGCTCTTACATTTTGG 1476
Db 38299 GCACTGTGGTTTTTAGAGTTTCAAGATTAACGTGCCAGTCTCTTCTAGCTTTTAGAG 38358
Qy 1477 TTTCTATTGGAAGTCAAGTGTGAATCTAATCAATCACTGCTTTTATATGTTAAATGCTCT 1536
Db 38359 TCTCTGTT-GAGAAATTTGAATCTGATAGTGTGAGCTTTTATATGTACCT-CTCC 38416
Qy 1537 TTTTCCCTGCACTTTTATATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1596
Db 38417 TCTTTCCCTCACAGCTTTTAAATATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 38476
Qy 1597 TATGCACTGTGGGAGTTTCTTTTCCGCTCAATCTATTTGGTGTGTTTGTATGCTTCTG 1656
Db 38477 TATATAGCAGGAGATTTCTTTTCTGCTCAGTCTA-ATGGTGTCTATAAGCTTTTG 38535
Qy 1657 TACCTGTATAGGCATCTCTTCTCAAGTTAGGAAATTTTCTTTTGGTTTTTCTTTGAA 1716
Db 38536 TATGTTTATAGGCACTCTTTCTTTAGTGTGGAAAGTTTCTTCTATGATTTTGTGGAG 38595
Qy 1717 AATATTTTCCCTGCTTTTGACCTG-----CCTTCTTCCCTTCCCTCTATTC- 1763
Db 38596 AGTATTTTCTGGCTTGGAGCTGGAGCTCTTCACTTCTTCTTCTTCTTCTTCTTCTTCT 38655
Qy 1764 -TTTGGTTTTGCAATGCTCTGGCTCTCTGATGTTTATGCTGCTGATTTATTTTGA 1822
Db 38656 GTTAGGCTTTTCAATGATCCAGATTTCTGATGTTTGTGTGAGGAAATTTTGTAGT 38715
Qy 1823 CTTAACATTTTCTTTGACCAAGTATCCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1882
Db 38716 TTTAACATTTTCTTTGACTGGTGA-----TCATCTATGCTGCTGAGAT 38757
Qy 1883 TCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1942
Db 38758 CCTCTCTTCCATCTCTGCTCTTCTGTTGGTGAATGCTTGGCTCTGTTGTTCTTCTTCT 38817

Qy 1943 CTTAAATTTTTCATTTCCAGATTTCTCAGTTTGGGTTTGGTTTATTATTACTTATTTCC 2002
Db 38818 CCTAGTTTTCACCTCCAG-----GGAAGTTTATGTTTCTTCTTATTACTTCTTATTTCC 38871
Qy 2003 ACTTTCCAGTCTGGAATGTTTACTCATTT---TTCTCCAGTATTTTACATTTTCATAG 2059
Db 38872 ACTTTCCAGTCTTACAGATTTTATTTATTTCTTCTCATCTGTTAATATTTTCTTAT 38931
Qy 2060 GTTCTTTTAAATGGAATTTATCTTCTCTTCAAGGACCTTTTATGAATTCATAAATGT 2119
Db 38932 AATCTTTTAAAGGAATTAATTTGTTTCTCTTTTAAAGACATCTACATGTTGATTTT 38991
RESULT 6
US-10-087-192-1663
; Sequence 1663, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1663
; LENGTH: 193853
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(193853)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1663

Query Match 26.8%; Score 610.3; DB 1; Length 193853;
Best Local Similarity 63.8%; Pred. No. 1.1;
Matches 1180; Conservative 0; Mismatches 552; Indels 117; Gaps 18;

Qy 292 TCTATTTCTGATTTCTATCTTGGCTCATTTTAACTCAGTAGTGTGTTGGTTTCC 351
Db 152240 TTTCTTTCTTTTATCTTCTTCCCTTGACCAAGGTATCATTTGAGAGTGTGTTTCAGTTTCC 152299
Qy 352 ATAAAGTTTGTAAATTTCTGTTTCTGTTTCTGTTTGTGTTTGTATCTAGATTAAGCTGT 411
Db 152300 ACGTGAATGTTGGCTTCTTATTTATTTTGTATTGAAGAT---CAGCCTTAGTCCAT 152356
Qy 412 GGTGTCAGATAGGACATAGATTTATTTCAATTTGTCCTTTTATCTGTCGAGACTTGTCT 471
Db 152357 GGTGATCTCATAGGATGTCATGGAACAAATTTCAATGTTTTGTATATGTTGAGGCTTGT 152416
Qy 472 TGTGTTTGAATATGATTTCAATTTTGGAGA--GTTTCATAGGTCGTGCAAGAGGTAC 529
Db 152417 TGTGACCAATTTATGTTGTCATTTTGGAGAGTACCATGAGTGTCTGAGAAGAGTAT 152476
Qy 530 AGTC-TTTCGTGTTTGGTGAATAGTCTGTAAATATCTCT-AGGTCCTAGTGGTTTATGA 587
Db 152477 ATCCCTTTTGTTTAGGATAAAATGTTCTGTAGATATCTGTCCAGATCCATTTGTTTCATA 152536
Qy 588 CATCAGTGTAGTCCAGCATTTCTGTTTCTGTTTCTGTTTGTGTGAGATGACCTACTGTGGA 647
Db 152537 CTTCTGTTAGTTTCACTGTGTCCTGTTTGTAG-TTTTGTTCACGATCTGCCATTTAGT 152595
Qy 648 GAGAATGGGTATTTGAAGTAGCCCACTATCTGTGTGTGAGGT-CAATATGTGATTTTACG 706
Db 152596 GAAAGTGTGTGTGTGAGTCTCCCACTATTTATGTTGAGGTGCAATGTGTCCTTTGAGC 152655

QY 707 TGTAGCTGTCTGTTTATGAACCTGGGTGACATTTGTTTGGTGCATAGACATTAAGA 766
Db 152656 TTACTAAAGTTCTTTAATGAATGTGGCTGCCCTGTTATTGAGCATAGAAATTCAGA 152715
QY 767 ATTGCAATGCTCTCTGGTGTGATTTTCCCTTGATGCTATGATTAATCTTCCCATCT 825
Db 152716 ATTGATAGTTCTCTGAGGATTTTACCTTTGATGAATATGAAGTGCCTCTCTGCT 152775
QY 826 CATCTGCTTAGTTTGGGTTTAAAGTCTA-TTAGT-CAGATATAAAATGACTGTATCGGCT 884
Db 152776 TTTTGTATGACTTTGGGTTGGAAGTCAATTTTATAGATATTAGATGGCTACTCCAGCT 152835
QY 885 TGCTTCTTAGGCCATTTGCTTAGAATA-TCTTTTCCATCTTTTACTCTTAAGTGATGT 943
Db 152836 TGTTTTTCATACATTTTGTCTGGAAATTTGTTTCCAGCCCTTTCAITCTGAGGTAGTGT 152895
QY 944 CTATCCAT--GGTAGGTTGCTTTTGGATGACAGTAGGATGGATCTGTTTTCAT 1000
Db 152896 CTAICTTTTCTCTGAGATGAGTTTCTGTAGAGCAAAATGTTGGTCTTGTGTGT 152955
QY 1001 ATCCATCTGTACCAGTATCTTTTCTAGAGAAATTAAGATCAATGAGTCATGATGT 1060
Db 152956 AGCCAGTTAGTTAGTCTATGCTGTTTATTTGGGAGTTGAG-----TCCATTTGATAT 153007
QY 1061 TGAGAAATTAATGAGCAGTGTGTTGATCTTGTATCTTGCACCTGTGAAGTGTGT 1120
Db 153008 TAAGAGATTAAGAAAGAAATGTTGCTTCTGTTATTTTCTTTTCTTTTCTTTAAG 153067
QY 1121 GT 1180
Db 153068 TTGGCAATCTGTCTTGTGGCTGCTCTTTTAGTTG----- 153106
QY 1181 GT 1240
Db 153107 -----TTGAGGGAT 153115
QY 1241 TTATTATTCAATTTTCTGAATGTGGTAAATCTTTAGATTAAGAAATTTTCTCCTAGC 1300
Db 153116 TACCTTCTGCTTTTCTAGGGCGTGGTTCGGTCTGTATGTTTCTTTTCTATTATT 153175
QY 1301 CTTCTTT--AGGTCTGCAATTTGAAGATAGATATTTCTTACATCTGATTTTATCTTGAAT 1358
Db 153176 ATCCCTTGAAGGGCTGGAATTTGTGAAAGATAATGTGTAATTTGGTTTGTCTGTGAAT 153235
QY 1359 GTCTTTCTTCTCCAACTATTGTGACAGAAATTTTCTAAGTCCAGTGTGCGCTGA 1418
Db 153236 ACTTTGGTTTCTCCATCTATGTAATGAGAGTTTGGCTGGGTATAGTACCTGGCTGG 153295
QY 1419 CATCTGTAGTCTCTGTGAGTCTGTAGCACAATCTGTGAGGGCTTCTTACATTTTGAATT 1478
Db 153296 AATTGTGTCTCTTGTAGTCTGTATACATCTGTCCAGGCTCTTCTGGCTTTCATAGTC 153355
QY 1479 TCTATTGAAAGAGTCAGGTAAATCTTAATACATCTGCTTTATATGTTAATGCTCTTT 1538
Db 153356 TCTGCT-GAAAAAATCTGTAATCTGATAGGCTGCGCTTATATGTTACTT-GACCTT 153413
QY 1539 TTTCCCTTGCATCTTTTAAATATCTTCTTGTGTTCTATATCTTTTAGTGAATTTGATTATTA 1598
Db 153414 TTTCCCTTACTGCTTTTAGTATCTATCTTTATTATAGTCAATTTGTTCTGTGATTATTA 153473
QY 1599 TGCATGTGGGAGTTCTTTTCCGGTCCAAATCTATTATGGTGTGTTTGTATGCTCTTGTGA 1658
Db 153474 TGTGTGGGAGGAATTTCTTTTCTGGTCCAGTCTATTTTGGAGTTCTGTAGGCTTCTTGTGA 153533
QY 1659 CTTGATAGGCAATCTTTTCAAGGTTAGGAATTTTCTTTTGTGTTTCTTGAATA 1718
Db 153534 TGTTCATAGGCAATCTTTTCTTTTAGATTGGAAGTTTCTTCAATATTTTGTGTAAGA 153593
QY 1719 TATTTTCCCTGTTTGTACCTG-----CCTTCTTCCCTTCTCTATTTCT 1764
Db 153594 TATTTGCGGGTCTTTGAGTTGAAATCTTCAATCTCTTCACTCTTATTTATCTGTAGGT 153653

QY 1765 TTGGTTTTTGCATAGTGTCTCTGGCTTCTGATGTTTTTATGCCCTGGATTTATTAGACT 1824
Db 153654 TTGGTCTTCTCATTGTTGCTCTGGATTTCTCTGGATTTTGTAGTTAGGATCTTTTTCATT 153713
QY 1825 TAACATTTTCTTTGACCAAGGTATCCATTTCTTATCTTGTCTTCACTGCTGAGATTC 1884
Db 153714 TTCCATTTTCTTTGATTTGTTGCGGATGTTCTCTATGGAATCTTCTGCACCTGAGATTC 153773
QY 1885 TCTCTTCTATCTCTGTATTTCTGTAGTGGCTGTCTCTGAGGTTCTGTG-TTGGGTTTC 1943
Db 153774 TCTCTTCCATCTCTGTATTTCTGTGTGCTGATCTCTCAAAATCTGTGTTCCAGATTTCTTTC 153833
QY 1944 TTAATTTTCTTCAATTTCCAGATTTCTCTAGTTTGGTTTGTGTTTATTAATCTTATTTCCA 2003
Db 153834 CTAGGTTTCTATCTCCAGCTTGTCTCACTTTGGGTTTCTTTATTTGTGCTACTTCTC 153893
QY 2004 CTTTCAGGCTCTGAAATGTTTACTCANTTCTCTCC---AGTATTTACATTTTCATAGG 2060
Db 153894 TTTTAGGTCTTGGATGGTTTATTTCAATCTCTCACTGTGTGTTGTTGTTTTCATGCA 153953
QY 2061 TTTCTTTAATGATTTATTCATTTCTCTTCAAGCACCTTTTATGAAT 2109
Db 153954 ATTCTTTAAGGATTTTGTGCTTCTCTTTTAAATGTCTTCTACTTGT 154002

RESULT 7
US-10-369-214-129/c
; Sequence 129, Application US/10369214
; Publication No. US20030232037A1
; GENERAL INFORMATION:
; APPLICANT: Groot, Pieter C.
; APPLICANT: Berghenhegouwen van, Bram J.
; APPLICANT: Coesterhout van, Antoon J.M.
; TITLE OF INVENTION: Genes involved in immune related responses observed
; TITLE OF INVENTION: with asthma
; FILE REFERENCE: P53837US00
; CURRENT APPLICATION NUMBER: US/10/369,214
; CURRENT FILING DATE: 2003-02-15
; PRIOR APPLICATION NUMBER: EP 00202867.8
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: PCT/NL01/00610
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 139
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 129
; LENGTH: 19619
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(19619)
; OTHER INFORMATION: /note="Contig 1A gene"

US-10-369-214-129

Query Match 26.0%; Score 591.499; DB 1; Length 19619;
Best Local Similarity 63.2%; Pred. No. 3.5; Indels 119; Gaps 16;
Matches 1138; Conservative 0; Mismatches 545

QY 292 TCTATTCTTGTATTTCTATCTTGGCTCATTTTAACTCAGTAGTGAGTTGTTGTTTCC 351
Db 19360 TTTCTTTCTTTATCTCTTCCCTTGACCAAGGTATCATTTGAGAAGAGTGTGTTTCACTTTCC 19301
QY 352 ATAGTTTGTAGTTTCTGTTGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 411
Db 19300 ATGTGAATTTGGCTTCTGTTTATTATTGTTTATTTGAAGAT---CAGCCTTAGTGAT 19244
QY 412 GGTGTTTCATAGGACATAGAGTATTATTCAATTTGCTTTTATCTGTCGAGACTTTGCTT 471
Db 19243 GGTGATCTGATAGGATACATGGCAATTTTCAATATTTTTGAATCTGTTGAGCCTGATT 19184
QY 472 TGTTTTGAATATGATTTCAATTTTGGAGA--GTTTCTAGGGTCTGACAGAGAGGTAC 529
Db 19183 TGTGACCTATTATGTTGTTCAATTTTGGAGAAGGTACCATGAGGTCTGAGAGAAGGTAT 19124


```

; FILE REFERENCE: 21199
; CURRENT APPLICATION NUMBER: US/10/388,934
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: 02005336.9
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 02015657.6
; PRIOR FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 862
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 580
; LENGTH: 6335
; TYPE: DNA
; ORGANISM: Rattus norvegicus (No. US20040005547Alway rat)
US-10-388-934-580

Query Match      25.8%; Score 588.5; DB 1; Length 6335;
Best Local Similarity 62.7%; Pred. No. 5.8;
Matches 1173; Conservative 0; Mismatches 560; Indels 137; Gaps 18;

QY 292 TCTATTCTTGATTCTATCTCTGCTCATTTTAACTCAGTAGAGTGTGTTTGTTTCC 351
Db 2595 TTTCCTTCTTATTCTTCCCTGACCAAGTTATCATAGTAGAGCAATGTTTCAGTTTCC 2536

QY 352 ATAAGTTTGAAGTTTCTGTGTTTCTGTTGTTTGTGTTTGTATCTAGATTTAAAGCTGT 411
Db 2535 AAGTATATATGGGCATTTC-----TTCCTTGATTGTTATTGAACACCAAGTTTAGGCCGT 2482

QY 412 GGTGGTCAGATAGACATAGAGTATTATTTCAATTGCTTTTATCTCTGTCAGACTTGCCT 471
Db 2481 GGTGGTCAGATAGATCATCGGGAATTATTTCTATCTTCTGTACCTGTTGAGGCCCGTTT 2422

QY 472 TGTTTTGAAATATGATATCAATTTTGGAG--AGTTTCATAGGGTGCTGCAAGAAGAGGTAC 529
Db 2421 TTTGACCAATTATATGCTCAATTTTGAGAAAGTACCATGAGGAGCTGAGAGAGGTAT 2362

QY 530 AGCTTTTGTGTTTGG-TGAAATAGTCTGTAAATAT-CTCTAGTCCACTTGTGTTTATGA 587
Db 2361 ATCCTTTTGGTTTAGGATAGAATGTTCTATAAATATCCGTTAAGTCCCTTTGGTCAATGA 2302

QY 588 CATCAGTTAGTCTCCAGCATTTCTCTGTTTCGTTTTGTTGAGATGACCTAACTGTTTGA 647
Db 2301 CTCTCTTAGTCTGTACATCTCTGTTT-AAITTCGTTTCCATGATCTGTCCATTGAT 2243

QY 648 GAGAAATGGGATATGAAGTAGCCACATCTGTGTGTAGGT-CAATATGTGATTTTTCAGC 706
Db 2242 GAGAGTGGGGTGTGAAATATCCACATATTATTTGTGTGAGGTGCAATGTGTGTTTTCAGC 2183

QY 707 TGTAGTCTGCTGTTTATGAATCTCGGTGCACATCTGTTTGGTGATAGACATTAAGA 766
Db 2182 TTTAGTAAGGTTTCTTTTACATGTAGGTGCCCTTGATTTGGGGCATAGATATTTAGG 2123

QY 767 ATTGCAATGCTCTCTGGTGA-TTTTTCCTTGTAGTGTATGTAGTATCTTCCCAATCT 825
Db 2122 ATTGAGAGTTCACTTGGTGGATTTTCCCTTGTATGAATATGAAGTGTCTCTCTTATCT 2063

QY 826 CATCTGCTTAGTTTGGGTTTAAGTCTA-TTAGTCAGATATTAATAATGACTGTATCGCT 884
Db 2062 TTTTGTAGTACTTTTAGTAAATAATGATTTATTTGATATTAAGATGGGTACTCCAGCT 2003

QY 885 TGCTTCTTAGGGCCATTGCTTAGAATA-TCTTTTCCATCCTTTTACTCTAAAGTGTATGT 943
Db 2002 TGCTTCTTCAGCAATTTGCTTGAAATATGTTTTCAGCCCTTTCATCTCTGAGGTATGT 1943

QY 944 CTATCCATG--GTAGGTTGCTTTTTTGGATGACAGTAGGATGAATCTTGTTTTCAT 1000
Db 1942 CTGCTTTGTCTCTGAGGTGTGTTTCTGTAGAGAAATAAGATCATTTGAGTCATTTGATGT 1060

QY 1001 ATCCATCTCTGTATCCCAAGTATCTTTTCTTAGAGAAATAAGATCATTTGAGTCATTTGATGT 1060
Db 1882 ATCCAGTTTGTAACTATGTGCTTTTATTTGGGAGTTAAG-----GCCATTTAATGT 1831

QY 1061 TGAGAAATTAATCAATGAGCAGTGTGTTGGAAATCTTGTATCTTGCACCTGTGAAGTGTGT 1120

```


Db 1126 TTACATTATCTTTGACAGTTCAGTCAATGATTTCTATGGAACTCTTCGCTCTTGAGATTC 1067
Qy 1885 TCTCTTCTATCTCTGTATTTCTGTCAGTGGCTGTCTCTGAGGTTC-TGTTGGGTTT 1943
Db 1066 TCTCTTCAATCTCTGTATTTCTGTGATGCTGTGATCTACGGCTCCCTGTCTCTTCC 1007
Qy 1944 TTAATTTTTCATTTCCAGATTTCTTCAGTTGGGTTTGTATTATTAATCTATTCCA 2003
Db 1006 TTGGTTTCTATATCCAGGCTTTTCCATGTCTCTTCTGATGCTTCATTTCCA 947
Qy 2004 CTTTCAGTCTCTGAATGTTTACTCATTTTCTCCAGATTTTACATTTTTCATAGTTT 2063
Db 946 TTTTAAATCTTCACTGTT-----TGATGTGTTTCTCTGGAAT 905
Qy 2064 CTTTAAATGGATTTATCATTTCTCTTCAAGGACCTTTTATGAATTCATTAATATGTAT 2123
Db 904 CTTTCAGGGAATTTTGGCAATCTCTCTGTAGGCTTCTACTGTCTCTTAAGGAGTTCT 845
Qy 2124 TAAGGTCTCT 2133
Db 844 TCACGTCTTT 835

RESULT 12

US-10-191-803-720/c
; Sequence 720, Application US/10191803
; Publication No. US20040014040A1
; GENERAL INFORMATION:
; APPLICANT: MENDRICK, Donna
; APPLICANT: PORTER, Mark
; APPLICANT: JOHNSON, Kory
; APPLICANT: HIGGS, Brandon
; APPLICANT: CASTLE, Arthur
; APPLICANT: ELASHOFF, Michael
; TITLE OF INVENTION: Cardiotoxin Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5090US
; CURRENT APPLICATION NUMBER: US/10191.803
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/303,819
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,623
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: US 60/369,351
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: US 60/377,611
; PRIOR FILING DATE: 2002-05-06
; NUMBER OF SEQ ID NOS: 1140
; SOFTWARE: Patent In Ver..2.1
; SEQ ID NO 720
; LENGTH: 6335
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20040014040A1 X53581
US-10-191-803-720

Query Match 25.8%; Score 588.5; DB 1; Length 6335;
Best Local Similarity 62.7%; Pred. No. 5.8;
Matches 1173; Conservative 0; Mismatches 560; Indels 137; Gaps 18;

Qy 292 TCTATTTCTCATTTCTATCTTGGCTCATTTTAACTCAGTAGTGAGTGTGTTGGTTTCC 351
Db 2595 TTTCTTTCTTTATTTCTTCTTGACAGGTTATCATTTAGTAGGCAATTTTCAATTTC 2636
Qy 352 ATAAAGTTTGAAGTTTCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 411
Db 2635 AAGTATATATGGCATTCT-----TTCCTTGATTGTTTATTGAAGCAGTTTATAGGCCGT 2482
Qy 412 GGTGTCAGATAGGACATAGATATTTTCAATTTGTTTATCTCTCTCGAGACTTGTCT 471
Db 2481 GGTGTCGATAGCATGATGAGGATTTATTTCTATCTTCTGTACCTCTGTAGGCCCGTT 2422

Qy 472 TGTTTTGAAATATGTAATCAATTTTGGAG--AGTTTCATAGGTCCTGACAAAGAGTAC 529
Db 2421 TTTGACCAATATATAGTCAATTTTGGAGAAAGTACCAATGAGGAGCTGAGAGAGGTAT 2362
Qy 530 AGTCTTTGTGTTTGG-TGAAATAGTCTGTAATAAT-CTTAGGTCCACTTGGTTTATGA 587
Db 2361 ATCCCTTTGCTTTAGGATAGAATGTTCTATAAATATCCGTTAAGTCCCTTTGGCTCATGA 2302
Qy 588 CATCAGTTAGCTCCAGCAATTTCTCTGTTTCGTTTTCGTTTTCGATGACCTAACTGTTGGA 647
Db 2301 CTTCTCTTAGTCTGTCTACATCTCTGTTT-AAATTTCTGTTTCCATGATCTCTCCATTTAT 2243
Qy 648 GAGAAATGGGTTATTCAGTAGCCCACTATCTGTGTGAGGT-CAATATGTAATTTTACG 706
Db 2242 GAGAGTGGGTTTGAATCTCCCACTATTTGTGTGAGGTGCAATGTGTGTTTGGAGC 2183
Qy 707 TGTAGTGTGCTGTTTATGAACCTTGGTGACATTTGTTTGGTGCATAGACATTAAGA 766
Db 2182 TTTAGTAAGGTTTCTTTTACATATAGTGGTGCCTTGTATTGGGGCATAGATATTAGG 2123
Qy 767 ATTGCAATGTCCTCTTGGTGA-TTTTCTTTGATGCTATGTAGTATTTTCCCAATCT 825
Db 2122 ATTGAGAGTTCACTTGGTGGATTTTCTTTGATGATATGAAGTGTCTTCTTATCT 2063
Qy 826 CATCTGTAGTTTGGGTTTAAAGTCTA-TTAGTCAGATATATAAATGACITGATCGGCT 884
Db 2062 TTTTGTGATGACTTTTGTAAATAATTTTATTTGATATTTAGAAATGGCTACTCCAGCT 2003
Qy 885 TGCTTTCTTAGGSCCAATTTGCTTAGAATA-TCTTTTCCATCTTTTACTCTTAAGGTGATGT 943
Db 2002 TGCTTTCTCTGACCAATTTGCTTGGAAATTTGTTTCCAGCCTTTCACITCAGGTAATGT 1943
Qy 944 CTATCCATG---GTAGGTGTCTTTTGGATGAGCAGTAGGATGGAATCTTGTTCAT 1000
Db 1942 CTGTCTTTGTCTCGAGGTGTGTTTCTCTAGGACAGCAATGCAAGGTCTCTGTTGCGT 1883
Qy 1001 ATCCATCTGTTTACCAGTATCTTTTCTAGAGAAATTAAGATCATTCAGTCAATGATGT 1060
Db 1882 ATCCAGTTTGTAAATCTATGCTTTTATTTGGGGAGTTAAG-----GCCATTTATGT 1831
Qy 1061 TGAGAAATATCAATGAGCAGTGTTTGTGGATTTCTTGTATCTTGCACCTTGTGAAGTGTGT 1120
Db 1830 TGAGAGATATTAAGGAATAGATTTATGCTTCCGTTTATATTCATATTTGGATGTGAGG 1771
Qy 1121 GT 1180
Db 1770 TTATGTTTGTGTG----- 1758
Qy 1181 GTGTCTGTGTGTGTGTGTGTGTGTCTCTCCCTCTTTTGATTTTGGCCTGGAATPAT 1240
Db 1757 -----CTTTCAATCGCTTTGTTTGTGTCAGACCAT 1725
Qy 1241 TTATTTATCATATTTTCTTGAAATGAGGTAAACATCTTTAGATTTGAAGTTTTTCTCCTAGC 1300
Db 1724 TAGTTTCTGTCTTCTTAGGATAGCTTGCCTCTTATGTGCGCTTTTACCAATTTAT 1665
Qy 1301 CTTCTT--TAGGTCTGCATTTGAAGATAGATATCTTTTACATCTCATTTTATCTTAGAAT 1358
Db 1664 ATCCCTTTGAGTGGCTGGAATTTGTAGAAAGATATGTGTAATTTGGTTTGTCTAGGAT 1605
Qy 1359 GTCTTTCTTCTCCAACTATTGTGACAGAAAGTTTTTCTAAGTGCAGTAGTCTGCGCTGA 1418
Db 1604 ATCTTGGTTTCTCCATCTATGTTAATTCAGAGTTTTTGCAGGATACAGTAACTGGGCTGG 1545
Qy 1419 CATCTGTAGTCTCTTGGAGTCTGTAGACATCTGTGACGGCTTCTTACATTTTGAAT 1478
Db 1544 CATTTGTGTCTCTTAGGCTCTGTATGACATCTGTCCAGGATCTTCTGGCCTTCTATGTT 1485
Qy 1479 TCTATTGGAAGAGTCAGGTGTAATTTCTAATACATCTGCTCTTTATATGTTAATTTGCTCTTT 1538
Db 1484 TCT-GGCGAGAGTCTGGGTGATTTCTCATAGGTCTGCTTTATATGTTACTT-GACCTT 1427
Qy 1539 TTTCCCTGCACTCTTTTAAATATTTCTTTTGTCTCTACTTTAGTGTATTTGATTA 1598

Db 1426 TTTCCTTACTGCTTTTAAATATCTTCTTTATTTTGGCGTTGGTGTGACTATTA 1367
Qy 1599 TGCACTGCGGAGTTCTTTTCCGGTCCAAATCAATTTGTTGTTGTTGATGCTTTGTA 1658
Db 1366 TGTGACGGAGGTGTTCTTTTCTGGTCCCAATCAATTTGAGTTCTGTAGGCTTTGTA 1307
Qy 1659 CTTGATAGGATCTCTTTTCAAGGTTAGAAATTTTCTTTTGGTTTCTTTGTA 1718
Db 1306 TGCCTATGGGTATCTCTTTTATAGTTAGGAAGTTTCTTCTATGATTTGTTGA 1247
Qy 1719 TATTTTCCCTGCTTTTGACCTG-----CCTTCTCCCTTCTCTATTCCT 1764
Db 1246 CATTTACTGTCCTTTGAGCTGGAGCTTCCACTCTCTTATACCTATTAATCCTTAGGT 1187
Qy 1765 TTGTTTTTGCATAGTGTCTGCTTCCGATGTTTATGCTGGATTAATTTAGACT 1824
Db 1186 TTGATCTTCTCATGAGTCTGATTTCTGTATGTTTGGACCAAGTATTTTCCGCT 1127
Qy 1825 TAAATTTTCTTACCAAGATCAATTTCTTCTATCTTCTTCTACTGCTGAGATTC 1884
Db 1126 TTACATTAATCTTACAGATGAGTCAATGATTTCTATGGAACTTCTGCTCTTGAGATTC 1067
Qy 1885 TCTCTTCTATCTTGTATTTCTGAGTGGCTTGTCTCTGAGGTTCC-TGTTGGGTTTC 1943
Db 1066 TCTCTTCATCTCTGTTATTTCTGTTGGTATGTTGATCTACGGCTCTTGTCTCTTC 1007
Qy 1944 TTAATTTTCTTCAATTCAGATTTCTTCAAGTTTGGTTTGTATTAATTTCTATTTCCA 2003
Db 1006 TTGGTTTCTTATATCCAGGTTGTTTCCATGTTCTTCTTCTGATTTCTTATTTCCA 947
Qy 2004 CTTTCAGGCTCGAAATGTTTTACTCATTTTCTCCCAAGATTTTACATTTTCTATGTTT 2063
Db 946 TTTTAAATCTTCAACGTGT-----TGATTTGTTTCTCTGGAAAT 905
Qy 2064 CTTTAAATGATTAATTAATTTCTTCAAGGACCTTTTATGAATTCATAAAATGTAAT 2123
Db 904 CTTTCAGGATTTTGGGATTCCTCTGTTAGGCTTCTACTTGTCTCTAAGGAGTTCT 845
Qy 2124 TAAGTCTCT 2133
Db 844 TCACGTCCT 835

RESULT 13
US-10-374-077-207
; Sequence 207, Application US/10374077
; Publication No. US20040006779A1
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; Yu, Chang-En
; Oshima, Junko
; Mulligan, John T.
; Schellenberg, Gerald D.
; TITLE OF INVENTION: ANTIBODIES AGAINST GENE PRODUCTS RELATED TO
; WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed Intellectual Property Law Group
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/374,077
; FILING DATE: 25-Feb-2003
; CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Rosenman, Stephen
REGISTRATION NUMBER: 43,058
REFERENCE/DOCKET NUMBER: 100107.401D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 207:
SEQUENCE CHARACTERISTICS:
LENGTH: 29604 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 207:
US-10-374-077-207

Query Match 25.8%; Score 588.3; DB 1; Length 29604;
Best Local Similarity 62.9%; Pred. No. 3;
Matches 1170; Conservative 0; Mismatches 567; Indels 123; Gaps 18;
Qy 292 TCTATTTCTGATTTCTATCTTGGCTCATTTTAACTCACTAGTAGTAGTTGTTGTTCC 351
Db 23722 TTTCTTTCTTTATCTTCTTCCCTTGACCAAGGTATCATTGAGAAGAGTGTATTTCAGTTCC 23781
Qy 352 ATAAGTTTCTGAAATTTCTGTTTCTGTTCTGTTGTTGTTGTTATCTAGATTTAAAGCTGT 411
Db 23782 ACGTGAATGTTGGCTTTCCATTTATTTGTTGTTATGAGATC---AGCCTTAGGCCAT 23838
Qy 412 GGTGTCAGATAGGACATAGAGTATTTATTTCAATTTGTTCTTTTATCTGTCGAGACTGCTT 471
Db 23839 GGTGTCCTGATAGGATACATGGACAAATTTCAATATTTTGTATCTATTGAGGCTGTATT 23898
Qy 472 TGTTTTGAATATGTTTCAATTTTGGAGA---GTTTCATAGGTCCTGACAGAGGTAC 529
Db 23899 TGTGACCAATTTATGTTCAATTTTGGAGAAGTCCCGTGGAGTCTGAGAAGAGGTAT 23958
Qy 530 AGTC-TTTTGTGTTTGGTGAATAGTCTGTAATATCTCT-AGGTCCACTTGGTTTATCA 587
Db 23959 ATCCTTTTCTTTAGGATAAATGTTCTGTAGATATCTGTAGGTCCTTGTTCATTA 24018
Qy 588 CATCAGTTAGCTCCAGCAATTTCTGTTTCTGTTGTTGTTGTTGAGATGACCTAACTGTTGA 647
Db 24019 CTTCCTGTTAGTTTCACTGTGCTCCTGTTTAG-TTCTGTTTCCAGCATCTGCTTTGAA 24077
Qy 648 GAGATGGGATTTGAAGTAGCCCACTATCTGTGTGTAGGT-CAATATGTCATTTTAC 706
Db 24078 GAAAGTGGTGTGTGAGTCTCCACTATTTATGTTGAGGTGCAATGATGCTTTGAGC 24137
Qy 707 TGTAGCTGTGTTGTTTATGAACCTTGGGTGACATTTGTTTGGTGCATAGACATTAAGA 766
Db 24138 TTTACTAAAGTGTCTCTAATGAATGTTGGCTGCCCTTGCAATTTGGTGCCTAGATATT 24197
Qy 767 ATTGCAATGCTCTTGGTGGATTTT-CCTTTGATGCCCTATGTTAGTATTTCTTCCCAATCT 825
Db 24198 ATTGAGTGTCTCTTGGAGGATTTTACCTTTGATGAGTATGAAGTGTCCCTTGTCT 24257
Qy 826 CATCTGCTTAGTTTGGGTTTAAAGTCTA-TTAGTCAGATATTAAATGACTGTATCGGT 884
Db 24258 TTTTGTATAAATTTGGTGTGGAAGTCGATTTTATCCGATACTAAAATGGCTACTCCAGCT 24317
Qy 885 TGTCTTCTAGGCCCAATTTGCTTAGAATA-TCTTTTCCATCTCTTTTACTCTCAAGGTGATGT 943
Db 24318 TGTTTCTTCAGTCCCAATTTGCTTGGAAATTTGTTTCCAGCCCTTTTACTCTGAGGTAGTGT 24377
Qy 944 CTATCCAT---GGTAGGTGTCTTTTGGATGTCAGCAGTAGGATGGATCTTTGTTTTCAT 1000
Db 24378 CTGTCTTTTCCCTGAGATGGGTTTCTGTAGCAGCAGAAATGTTGGTCTCTGTTTGTGT 24437
Qy 1001 ATCCATCTGTTTACCAGTATCTTTTCTAGAGAAATTAAGATCATTTAGTCAATTTGATGT 1060
Db 24438 AGCCAGTCTGTAGTCTATGTTCTTTTATTTGGAATTTGAG-----TCCATTGATAT 24489
Qy 1061 TGAGAAATTAATCAATGAGCAGTGTGTTGTAATCTTTGTTATCTTGGCATTGTGAGGTGT 1120

```
Db 24490 TAAGAGATATTAAAGAAAGTAATGTGCTCTCTTTATTTTGTGTAGAGTTGGCA 24549
QY 1121 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1180
Db 24550 TTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 24582
QY 1181 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1240
Db 24593 -----TTGAATGAT 24591
QY 1241 TTATTAATCATATTTCTTGAATGTGGTAACATCTTTAGATTTGAAGTTTCTCTCCAGC 1300
Db 24592 TACTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 24651
QY 1301 CTCTCTTT--AGGTCTGCATTTGAAGATAGATATTTTACATCTGATTTTATCTTAGAAT 1358
Db 24652 ATCTTTGAAGGGCTGGATTCGTGAAAGATATTTGTGTAATTTGTGTGTGTGTGTGT 24711
QY 1359 GTCCTTTCTCTCAACTATTTGTGACAGAAAGTTTCTTAAGTGCAGTGTCTGGCTGA 1418
Db 24712 ACTTTGGTTCTCCATCTATGTAATTTAGAGATTTGGCTGTATAGTCTGGCTGG 24771
QY 1419 CATCTGTAGTCTCTTTGGAGTCTGTAGCACATCTCTGTGAGGGCTCTTACATTTTGAGTT 1478
Db 24772 CATTTGTGTCTCTTAGTTCTGTATTAACATCTGTCCAGGCTCTTCTGGCTTTCATAGTC 24831
QY 1479 TCTATTTGAAAAGTCAGGTGTAATCTTAATACATCTGCTTTATATGTTAAATGCTTTT 1538
Db 24832 TCTGCT--GAAAAGTCTGGTGTAAATCTGTATAGGCTTCTTTATATGTTACTT--GACCTT 24889
QY 1539 TTCCCTTGATCTTTTAAATCTCTTTCTTTGTCTATATCTTTTGTAGTATTTGATTA 1598
Db 24890 TCTCCCTTACTGCTTTTAAATCTTAATCTTTTATTTAGTGAATTTGTGTCTGATTA 24949
QY 1599 TGCACCTGTGGGAGTTCTTTTCCGGTCCAACTATTTTGTGTGTGTGTGTGTGTGTGT 1658
Db 24950 TGTGTGGGAGGAATTTCTTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 25009
QY 1659 CTTGTATAGGATCTCTTTCTCAAGGTGTAGGAATTTTCTTTTGTGTGTGTGTGTGTGT 1718
Db 25010 TGATCAAGGATCTCTTTTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 25069
QY 1719 TATTTTCCCTGTTTGTACCTG--CTTCTTCCCTTCCCTCTATTCCT 1764
Db 25070 TATTAGCTGGCCCTTAAAGTTGAAATCTTCAATCTCATCAATCTCTATTCGTAGT 25129
QY 1765 TTGGTTTTGATAGTGTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1824
Db 25130 TTGCTCTCTCATTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 25189
QY 1825 TAAATTTTCTTTGACCAAGGTATCCATTTCTTATCTTGTCTTCTACTGCTGCTGAGATTC 1884
Db 25190 TTGTATTTTCTTTGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 25249
QY 1885 TCTCTTATCTTGTATTTCTGTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1943
Db 25250 TCTCTTCAATTTCTTGTATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 25309
QY 1944 TTAATTTTCTCAATTTCCAGATTTCTCTCAGTTTGGTTTGTGTGTGTGTGTGTGTGTGT 2003
Db 25310 CTAGATTTCTATCTCCAGGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 25369
QY 2004 CTTTCAAGTCTCGAAGTTTATCTATTTTCTTCCCGAGTATTT--TACATTTTCAAGG 2060
Db 25370 CTTTATGTTCTAGTATGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 25429
QY 2061 TTTCTTTAATGATTTATTTATTTCTTCCCTTCCAGGACCTTTATGAATTCATAAAGTA 2120
Db 25430 TTTCTTTAATGATTTCTACCTGTTTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 25489
```

```
US-10-322-281-115
; Sequence 115, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 115
; LENGTH: 289190
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(289190)
; OTHER INFORMATION: n = A,T,C or G
US-10-322-281-115
```

```
Query Match 25.8%; Score 587.8; DB 1; Length 289190;
Best Local Similarity 62.4%; Pred. No. 1.1;
Matches 1168; Conservative 0; Mismatches 582; Indels 122; Gaps 17;

QY 292 TCTATTTCTTGATTTCTATCTTTGGCTCATTTTAACTCAGTAGTGAGTTGTTGTTTCC 351
Db 132550 TTTCTTTCTTTATTTCTTCTTGTGACCAAGGTATCATTTGAGAAGTGTATTTCACTTCC 132609
QY 352 ATAAAGTTGTAAAGTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 411
Db 132610 ACCTGAAATGTTGGCTTCTCCATTTATTTATGTTGTATTGAAGATC---AGTCTTAGCCAT 132666
QY 412 GGTGTCTCAGATAGGACATAGAGTATTTTCAATTTGTCTTTTATCTGTCGAGACTTGTCT 471
Db 132667 GGTGTCTGATAGGATACATGGCAATTTCAATATTTTGTATCTATAGGCTGTCT 132726
QY 472 TGTTTTGAATATGTATTAATTTTGGAGA--GTTTCTAGGGTCTGACAGAGGTAC 529
Db 132727 TGTGACCAATATATATGTTCAATTTTGGAGAAGTCCCGTGGAGTCTGAGAAGAGGTAT 132786
QY 530 AGTC--TTTGTGTGTGTGAAATAGTCTGTAATATCTCT--AGTCCACTTGGTTTATGA 587
Db 132787 ATCTCTTTTGTGTGATAAAATGTTCTGTAGATATCTGTGAGTCCATTTGTTTCATAA 132846
QY 588 CATCAGTTAGCTCCAGCATTTCTCTGTTTCTGTTTGTGTGTGTGTGTGTGTGTGTGTGTGGA 647
Db 132847 CTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGAA 132905
QY 648 GAGATGGGGTATTTGAAGTAGCCCACTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 706
Db 132906 GAAAGTGTGTGTGTGAAAGTCTCCCACTATTTATTTGTGAGGTGCAATGTATGCTTTGAGC 132965
QY 707 TGTAGCTGTGCTTTTATGAACTTGGTGACATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGGA 766
Db 132966 TTTTACTAAGTGTCTTAAATGAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 133025
QY 767 ATTGCAATGCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 825
Db 133026 ATTGAGAGTCTCTTGGAGGATTTTACCTTTGATGAGTATGAAGTGTGTGTGTGTGTGTGT 133085
QY 826 CATCTGCTTAGTTTGGGTT--TAACTTATTAAGTATGAATTTAAATGACTGTATCGGC 883
Db 133086 TTTTGTGATGAGTTGGGTTGGAAGTCCGATTTTATCCGATATTAATATGAGTGTGTGTGT 133145
QY 884 TTGCTTCTTAGGCTCATTTGCTTAGAATA--TCTTTTCCATCTTTTACTCTAAGGTGATG 942
Db 133146 TTGTTCTTTCAGTCCATTTGCTTGGAAATTTGTTTCCAGCTTTTCACTCTGAGGTAGT 133205
QY 943 TCTATCCAT----GGTAGGTGTCTTTTGTGATGACAGTAGATGATGATGATGATGATGATG 998
Db 133206 TCTGTCTTTTCCCTGAGATGGGTTTCTCTGTAAGACAGAAATGTTGGTCTCTGTGTGT 133265
```

QY 999 ATATCCATCTGTATCCAGATATCTTTTCTAGAGAAATTAAGATCATAGTCAATGAT 1058
Db 133266 GTAGCAGTCTGTAGTCTATGCTCTTTTATTTGGGAATTGAG-----TCCATGAT 133317
QY 1059 GTTGAGAATTATCAATGAGCAGTGTTCGTGATCTCTGATCTTGCACCTGTGAAGTGT 1118
Db 133318 ATTAGAGATATTAAGAAAGATTAATGCTCTCCCTTTATTTTGTGTAGAGTTGG 133377
QY 1119 GT 1178
Db 133378 CATCTGTCTGT 133412
QY 1179 CTGT 1238
Db 133413 -----TTGAATG 133419
QY 1239 ATTTATTTATCATATTTTCTTGAATGTGGTAACATCTTTAGATTTGAAGTTTTCTCTTA 1298
Db 133420 ATTACTTTCTTGGTTGTCTAGGCGTGATTCGTCCTGTATGTCTCTTTCTCTTA 133479
QY 1299 GCCTTCTTT--AGGTCGCAATTTGAAGATAGATATCTTTACATCTCAATTTATCTTGA 1356
Db 133480 TTATCTTTGAAGGCTGTGATTCGTGGAAAGATATTTGTGAACCTTGGTTTTGTCTGGA 133539
QY 1357 ATGTCTTTCTTTCTCCAACTATTTGTGACAGAAAGTTTTTCTAAGTCAGTAGTCTGSCCT 1416
Db 133540 ATACTTTGGTTCTCCATCTATGTAATTTAGAGTTTGGCGGTAATAGAGCTGGCT 133599
QY 1417 GACATCTGTAGTCTCTTGGAGTCTGTAGACATCTGTGACGGGCTCTTACATTTTGA 1476
Db 133600 GGCATTTGTGTCTCTTAGTCTGTATTAACATCTGTCCAGGCTCTTCTGGCTTCATAG 133659
QY 1477 TTTCTATTGAAAGTCAAGGTGAATCTAATACATCTGCTTTATATGTTATGTTGCT 1536
Db 133660 TCTCTGTT--GAAAGTCTGTGTATTTCTAGAGGCTCTCTTTATATGTTTACTT--GACC 133717
QY 1537 TTTTCCCTTGCATCTTTAATATCTTTCTTTCTTCTATCTATCTTTTAGTATTTGATTTAT 1596
Db 133718 TTTCTCCCTTACTCTTTAATATCTTTATCTATCTTTATTTAGTCAATTTGTTCTCTGATTTAT 133777
QY 1597 TAGCATCTGGGAGATTTCTTTTCCGTCGCAATCTATTTGGTGTGTGTGTGTGTGTGT 1656
Db 133778 TATGTGTGGGAGAAATTTCTTTTCTGTCCAGTCTATTTGGAGTCTCTGTATGCTTCTTG 133837
QY 1657 TACCTTGATAGGCATCTTTCTCAAGTTAGGAAATTTTCTTTTGTGTGTGTGTGTGTGT 1716
Db 133838 TATGATCATGGGATCTCTTTTATGTTTGGAGAGTTTCTTCTATATTTTGTGTGTGT 133897
QY 1717 AATATTTTCCCTGTTTGAACCTG-----CCTTCTCCCTTCCCTCTATTC 1762
Db 133898 GATATTAGCTGGCCCTTTAAGTTGAAATCTTCAATCTCATCAATTCCTATTATCCGTAG 133957
QY 1763 CTTTGGTTTTTGCATAGTCTCTGGCTTCCGTGATGTTTATATCCCTGGATTTATTTAGA 1822
Db 133958 GTTTGGTCTCTCATGTGTCTCTGATTTACCTGGATGTTTGTAGTTAGGATCCCTTTTGA 134017
QY 1823 CTTAACAATTTCTTTGACCAAGGATCCATTTCTCTATCTTGTCTTCACTGTGCTGAGAT 1882
Db 134018 TTTTGTATTTCTTTGACTGTGTGTGCGATGTTCTCTATGGAATCTTCTGCACTGAGAT 134077
QY 1883 TCTCTCTCTATCTCTGTATCTGTCTGAGGAGCTGTCTGAGGTTCTCG--TTGGGT 1941
Db 134078 TCTCTCTCCATTTCTTTGTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 134137
QY 1942 TCTTAAATTTTTCATTTCCAGATTTCTCTCAGTTTGGTGTGTGTGTGTGTGTGTGTGTGT 2001
Db 134138 TCTTAGGTTTCTATCTCCAGGTTGCTCGTTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 134197
QY 2002 CACTTTCAGGTCGAAATGTTTACTCATTTTCTTCCCTCCAGATATTTACATTTTCAATAGGT 2061
Db 134198 CCTTTTGTAGTCTAGT 134257

QY 2062 TTCTTTAATGGAATTTATTCATTTCTCTCAAGGACCTTTTATGAATTCATAAAATGTAT 2121
Db 134258 CTTTCTTTAAGAGCCGTGAACCTCTTTAGCAGTCTCTCTCTGTAAATCTTTTAGTGACTT 134317
QY 2122 GTTAAAGTCTCTT 2133
Db 134318 ATGAAAGTCTCTT 134329

RESULT 15

US-10-175-523-57/c
; Sequence 57, Application US/10175523
; Publication No. US20030096264A1
; GENERAL INFORMATION:
; APPLICANT: Brockman, Jeffrey
; APPLICANT: Evans, David
; APPLICANT: Hook, Derek
; APPLICANT: Klimczak, Leszek
; APPLICANT: Laeng, Pascal
; APPLICANT: Palfreyman, Michael
; APPLICANT: Rajan, Prithi
; TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)
; FILE REFERENCE: 3235/10795-US3
; CURRENT APPLICATION NUMBER: US/10/175,523
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 60/299,151
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/317,828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/325,150
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/333,047
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 60/349,936
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/361,834
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57
; LENGTH: 225883
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-175-523-57

Query Match 25.7%; Score 585.399; DB 1; Length 225883;
Best Local Similarity 63.6%; Pred. No. 1.2; Indels 120; Gaps 17;
Matches 1131; Conservative 0; Mismatches 526;
QY 292 TCTATTTCTGTATTTCTATCTTTGGCTCATTTTAACTCAGTAGTGTGTTTGTGTTTCC 351
Db 187754 TTTCTTTCTTTATCTCTTCCCTTGACCAAGGTATCATTTGAGAAGAGTGTGTTTTCAGTTTCC 187695
QY 352 ATAAAGTTTCTGAAGTTTCTGTTTCTGTTTCTGTTTGTGTTGTATCTAGATTTAAGCTGT 411
Db 187694 ATGTGAATGTTGGCTTTTCCATTTATTTATGTTGTTTATTAAGAT--CAGTCTTAGGCCAT 187638
QY 412 GGTGGTCAGATAGGACATAGATATTTTCAATTTGTCTTTTATCTGTGAGACTTGTCTT 471
Db 187637 GGTGGTCTGATAGGATACATGGACAATTTCAATTTTGTATCTGTTGAGCCCTGTTT 187578
QY 472 TGTTTTGAATAATGATTTCAATTTTGGAGA--GTTTCAATAGGCTGTGACAGAAAGGTAC 529
Db 187577 TGTGACCAATTTATAGTTCATTTTGGAGAAGTCCCATGAGTGTGCTGAGAAGAGGTAT 187518
QY 530 AGTC-TTTCGTGTTTGGTGAAATAGTCTGTAAATATCTCT-AGGTCCTACTTGTGTTATGA 587
Db 187517 ATCCCTTTGTTTGTAGATAAATGTTCTGTAGATATCTGTGAGGCCATTTTGTTCATAA 187458
QY 588 CATCAGTTAGTCCAGCAATTTCTCTGTTTGTGTTTGTGAGATGACCTAACTGTTGGA 647
Db 187457 CTTCTGTTAGTTCATCTGTGTCCCTGTTTAG-TTCTGTGTTTCCAGATCTGTCCATTCAT 187399

QY 648 GAGAAATGGGGTATTGAAGTACCCCACTATCTGTGTGAGGT-CAATATATGATTTTACG 706
Db 187398 GAAAGTGGTGTGTTGAAGTCTCCCACTATTATTGTGTGAGGTGCAATGTGTGTTGAGC 187339
QY 707 TGTAGCTGTGCTGTTTATGAACCTTGGGTGACATCTGTTGGTGCATAGACATTAAGA 766
Db 187338 TTTACTAAAGTGTCTTAAATGAATGTGGCTGCCCTTGCATTTGAGGCTAGATATTGAGA 187279
QY 767 ATTGCAATGTCTCTTGGTGTGATTTT-CCTTTGATGCTTATGTAGTATTTCTTCCCAATCT 825
Db 187278 ATTGAGAGTTCATCTGGAGGATTTTACCCTTGAAGTATGAAGTGTCCCTCTCTGTCT 187219
QY 826 CATCTGCTAGTTTGGTGTAAAGTCTA-TTATGTCAGATATTAATAGACATCTGCGCT 884
Db 187218 TTTTGTAACTTTGGGTGAGGTGCAATTTATCCGATATTAAATGGGTCTCCAGCT 187159
QY 885 TGCTTCTTGGGCCATTTGCTTGAATA-TCTTTTCCATCTTTTACTCTAAGGTGATGT 943
Db 187158 TGTCTTTCAGATCATTTGCTTGGAAATTTGTTTCCAGCTTTTACTCTGAGGTAGTGT 187099
QY 944 CTATCCAT---GGTAGTGTCTTTTGGGANGCAGCAGTAGGANGATCTTGTTCAT 1000
Db 187098 CTGTCTTTTCCCTGAGATGGTTCCTGTAAAGCAGCAGAAATGTGTCTCTGTGTGT 187039
QY 1001 ATCCATCTGTACCCAGTATCTTTTCTAGAGAAATTAAGATCATTTGAGTCAATGATGT 1060
Db 187038 AGCCAGTCTGTATGTCTATGTCTTTTATTTGGGAATGAG-----TCCATGTAT 186987
QY 1061 TCAGAAATPATCAATGACAGAGTGTGTTGGATCTTGTATCTTGCACTGTGAAGTGTGT 1120
Db 186986 TAAGAGATTAATAGAAAGTAATTTGTGCATCTTTTATTTTGTGTGTACAGTTGCA 186927
QY 1121 GT 1180
Db 186926 TTCTGTCTTGT 186894
QY 1181 GT 1240
Db 186893 -----TTGAGTGTAT 186885
QY 1241 TTATATTCATATTTCTTGAATGTGGTAAATCATCTTTAGATTAAGATTTTCTCTAGC 1300
Db 186884 TACTTCTGT 186825
QY 1301 CTCTCTT---AGTCTGCAATTTGAAGATAGATATCTTTACATCTGATTTTATCTTGAAT 1358
Db 186824 ATCTTTTCATGGCTGGATTTCTTGAAGATATCGTGTGAATTTGGTTTGTCTGTGAAT 186765
QY 1359 GTCTTTCTTCCCACTATTGTGACAGAAAGTTTCTTAAGTGCAGTAGTCTGGCCTGA 1418
Db 186764 ACTTTGGTTTCTACATCTAGTAAATGAGAGTTTGGCCGGTATAGTAGCCTGGGCTGG 186705
QY 1419 CATCTGTAGTCTCTTGGAGTCTGTAGACATCTGTGAGGCCCTCTTACATTTTGATTT 1478
Db 186704 CATTTGTCTCTCTTGT 186645
QY 1479 TCTATTGAAAAGTCAGGTGTAATTTCTAATACATCTGCTTTATATGTTAATTTGCTTT 1538
Db 186644 TCTGTG-GAAGTCTGGTGAATTTCTGATAGGCCCTTCTTTATATGTTACTT-GACCTT 186587
QY 1539 TTTCCCTTGCATCTTTTAATATCTTTCTTGTGTTCTATACATTTTGTAGTATTTGATTA 1598
Db 186586 TCTCCCTTACTGCTTTAATATCTTATCTTTTATTTAGTGCATTTGTTGTCTGATTA 186527
QY 1599 TGCATGTGGGAGTCTTTTCCGGTCCAATCTATTGTGTGTGTGTGTGTGTGTGTGTGT 1658
Db 186526 TGTGTGGGTGGAATTTCTTTTCTGGTCCAGTCTATTGTGAGTCTCTGTAGGCTTGTGA 186467
QY 1659 CTTGATAGGATCTCTTTCTCAAGGTTAGAAATTTTCTTTTCTTTTCTTTTCTTTGAAA 1718
Db 186466 TGATCATGGCATCTCTTTCTTTATGTTTGGGAAAGTTTCTTCTATTATTATTGTGA 186407
QY 1719 TATTTTCCCTTGTTTGTGACCTG-----CCTTCTTCCCTCTCTATTCTT 1764

Db 186406 TATTAGCTGGCCCTTTAAGTTGAAAATCTTCATTCATCAATTCCTATTATCTGTAGGT 186347
QY 1765 TTGGTTTTTGTAGTAGTGTCTCTGGCTTCTCTGGAATTTTTATGCCCTGGATTTTAGACT 1824
Db 186346 TTGGCTTCTCTCATTGTGTCTGGAATTAACCTGGAATTTTGGATTTAGGATCTTTTGCA 186287
QY 1825 TAAATTTTCTTTGACCAAGGATTCATTTCTTATCTGTCTTCACTGCTGAGATTC 1884
Db 186286 TTGTATTTTCTTTGACGTGTGTGTGCAATTTCTCTATGAAATCTTCTGCACTGAGATTC 186227
QY 1885 TCTCTTCTATCTTGTGTATCTCTGAGTGCAGCTGTCTCTGAGGTCTCTG-TTGGGTTC 1943
Db 186226 TCTCTTCATTTCTTGTATCTGTTGCTGTGATGCTGCAATCTATGTTCCAGATCTTTC 186167
QY 1944 TTAATTTTCTTATTTCCAGATTTCTTCAAGTTTGGGTTTTGTTTATTAATTTCTATTCCA 2003
Db 186166 CTAGGTTTCTATCTCCAGCGTTGCTTGTCTTGGGTTTTCTTTATTTGTCTACTTCCC 186107
QY 2004 CTTTCAGGCTCGAATGTTTACTCATTTTCTCTCCC 2040
Db 186106 CTTTAGGCTAGTATGTTTGTTCATTTCCATCAC 186070

RESULT 16
US-10-003-806-10/c
; Sequence 10, Application US/10003806
; Publication No. US20020119229A1
; GENERAL INFORMATION:
; APPLICANT: Bishop, Colin E.
; APPLICANT: Agoulnik, Alexander I.
; APPLICANT: Zhu, Qichao
; TITLE OF INVENTION: CAN 1 AND ITS ROLE IN MAMMALIAN INFERTILITY
; FILE REFERENCE: P02066US1/10024824
; CURRENT APPLICATION NUMBER: US/10/003,806
; PRIOR FILING DATE: 2001-11-02
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; TYPE: DNA
; ORGANISM: Mouse
US-10-003-806-10

Query Match 25.7%; Score 585.1; DB 1; Length 173808;
Best Local Similarity 62.8%; Pred No. 1.4;
Matches 1188; Conservative 0; Mismatches 569; Indels 123; Gaps 18;
QY 292 TCTATTCTTGTATTTCTATCTTTGGCTCATTTTAACTCAGTAGTGAGTTGTTGTTTCC 351
Db 123632 TTTCCTTTCTTTATCTCTTCCCTTGACCAAGGTATCATTCAGAAGAGTGTGTTTCACTTCC 123573
QY 352 ATAAAGTTTGAAGTTTCTGTTGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 411
Db 123572 ATGTGAAGTTTGGCTTCTGTTTATTTATTTTGTATTGAAGAT---CAGCCTTAGTGGAT 123516
QY 412 GTGTGTTCAGATAGGACATAGAGTATTTTCAATTTGCTTTTATCTGTGAGACTTGTGT 471
Db 123515 GTGTGTCTGATAGGATACATGAGCAATTTCAATATTTTGAATCTGTTAGGCCCTGAT 123456
QY 472 TGTTTTGAATATGATTAATCAATTTTGGAGA--GTTTCATGGTCTCAGACAGAGTAC 529
Db 123455 TGTGACCTAATTATGTTGGTCAATTTTGGAGAAGGTACCAGGTGCTGAGAGAAGGTAT 123396
QY 530 AGTC-TTTCTGTTTGGTGAAATAGTCTGTAATATCTCT-AGGTCCACTTGGTTTATGA 587
Db 123395 ATCCCTTTGTTTGGTGAATATGTTCTGTAGATATCTGTAGATCTCTGAGATCTTTCATCA 123336
QY 588 CATCAGTTAGCTCCAGCATTTCTCTGTTTGTGTTTGTGAGATGACCTAACTGTTGA 647
Db 123335 CTTCTGTTAGTTTCAAGTGTGTCCTGTTTGTAGTTTCT-GTTTCCATGATCTGTCCTAGT 123277

QY 1719 TATTTTCCCTGCTTTTGACCTG-----CCTTCTCCCTTCTCTTATCT 1764
Db 122284 TGTTCCTGGACCTTTGAGTTGAAATCTTCACTCATCCACTCTATTTATCCGTAAGT 122225
QY 1765 TTGGTTTTGCAATAGTGTCTCTGGCTTCTGGATGTTTTATGCTCGATTTATTTAGACT 1824
Db 122224 TTGGCTTTCTTATTTGCTGCTGATTTCTGATATTTTGGATTTAGTATTTTGCATT 122165
QY 1825 TAACATTTCTTTTGACCAAGGATTCATTCTTCTATCTTTCTTCACTGCTCGAGATTC 1884
Db 122164 TTCCATTTCTTTGATTTGTCGCGATGTTCTCTATGGAATCTCTGACCTGAGATTC 122105
QY 1885 TCTCTTCTATCTCTTATTTCTGCTAGTGAAGCTTCTCTGAGGTTCTCTG-TTGGGTTT 1943
Db 122104 TCTTCTCACTCTCTTATTTCTGTTGCTGATGCTCAATCTATGTTTCCAGATTTCTTTT 122045
QY 1944 TTAATTTTTCATTTCCAGATTTCTTCAAGTTTGGGTTTTTGTATTAATTTCAATTTCCA 2003
Db 122044 CTAGGCTTTCTATCTCTAGTGTGCTCGCTTTGAGTTTTCTTTATTTGTTCTACTTCCC 121985
QY 2004 CTTTCAAGTCTCGAATGTTTTACTCAATTTTCTCTCC-----AGTATTTACATTTTCATAG 2060
Db 121984 TTTTATAGTCTAGTATGTTTGTTCATTTCCATTCACCTGTTTGTATGTTTTTCTCTCT 121925
QY 2061 TTTCTTTTAAATGGAATTTATTTCAATTTCTCTTCAAGGACCTTTTATGAATTCATAAAATGTA 2120
Db 121924 TTTCTGAAGACTTCTACTCTGTTTGAATGTTGTTTCTTCTTTCTTTAAGGACTTTGA 121865

RESULT 17
US-10-322-281-271
; Sequence 271, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 271
; LENGTH: 120239
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(120239)
; OTHER INFORMATION: n = A,T,C or G
US-10-322-281-271

Query Match 25.5%; Score 580.301; DB 1; Length 120239;
Best Local Similarity 61.9%; Pred. No. 1.7;
Matches 1199; Conservative 0; Mismatches 602; Indels 135; Gaps 19;

QY 292 TCTATTTCTTGAATTTCTATCTTGGCTCAATTTTAACTCAGTAGTAGTGTGTTGTTCC 351
Db 21580 TTTCTTTCTTTATTTCTCTTCCCTGACCGAGGTATCATTTGAGAAGAGTGTGTTTCAGTTTCC 21639
QY 352 ATAGTTTCTAGTTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTAGATTTAAGCTGT 411
Db 21640 ACCTAAATGTTGACTTTCTTATTTATTTATTTGTTTATTTGAAGATC---AGCTTAGTCCAT 21696
QY 412 GTTGGTTCAGATAGGACATAGATTTATTTCAATTTGTCTTTTATCTGTCTGAGACTTGTCT 471
Db 21697 GGTGATCTGATAGGATGATGAGCAATTTCAATATTTTGTATCTGTTAAGCCCTGTTT 21756
QY 472 TGTTTGAAATATGATTTCAATTTTGGAGA--GTTTCAAGGTCGTGACAAGAGGTAC 529
Db 21757 TGTGACCAAGTATATGATCAATTTTGGAGAAGGTAGCATGAGTGTGATGAGAAGAGGTAT 21816
QY 530 AGTCTTTGTTGTTTGTGTAATAGTCTGTAAATATCT-CTAGTCCACTTGGTTTATGAC 588

QY 648 GAGAAATGGGTAATTTGAAGTAGCCACATCTATCTGTGTGAGGT-CAATATGATGATTTTACG 706
Db 123276 GAAAGTGTGTGTTGAAAGTCTCCCACTATTATTGTTGAGGTGCAATGTTGTCTTTGAGC 123217
QY 707 TGTAGCTGTGCTGTTTATGCACTTGGTGACATTTGTTTGTGATGATGACATTAAGA 766
Db 123216 TTTACTAAAGTTTCTTTAGTGAATGTTGCTCTCTTTGATTTGGAGCATAGATTCAGA 123157
QY 767 ATTGCAATGCTCTTGTGATTTT-CCTTTGAAGCTATGATAGTATTTCTTCCCAATCT 825
Db 123156 ATTGAGAGTTCTCTTGGAGGATTTTACCTTTGATGAGATGAAGTGTCCCTCTCTGTCT 123097
QY 826 CATCTGCTTAGTTTGGGTTTAAAGTCTAT-TAGTCAGATATTAATAGCTATCGCT 884
Db 123096 TTTTGTATGACTTTGGGTTGGAAGTCAATCTTATCAGATATTAGATGGCTACTCTCTGCT 123037
QY 885 TGCTTTCTTAGGGCCATTTGCTTTAGAATA-TCTTTTCCATCTCTTTTACTCTAAAGGTGATGT 943
Db 123036 TGTTCITCATACCATTGCTTTGGAATTTGTTTCCAGCTTTTCAITCTCAGGTAGTGT 122977
QY 944 CTATCCAT---GGTAGTTGCTTTTGGATGACAGATGAGATCTGTTTTCAT 1000
Db 122976 CTATCTTTTCTCTCAGATGATTTTCTCTGTAAGCAGCAAAATGTTGGGCTCTGTTTGTGT 122917
QY 1001 ATCCATTTCTGTACCCAGTATCTTTTCTAGAGAAATTAAGATCATTCAGTCTATGATGT 1060
Db 122916 AGCCAGTTTGTAGTCTATGCTTTTATTTGCGAGTTGAGA-----CCATGATGT 122865
QY 1061 TGAGAAATATCAATCAGCAGTGTTTGTGGATCTTGTATCTTCACTTGTGAAAGTGTGT 1120
Db 122864 TAAGAGATATTAAAGCAAAAGTAATTTGCTCTCTGTTATTTTATTTAGTTTAAAGGTGCA 122805
QY 1121 GT 1180
Db 122804 TTTGTTTCTTGTGGCTGTCTTCTTTAGTTTG----- 122772
QY 1181 GT 1240
Db 122771 -----TTGAGGAT 122763
QY 1241 TTATTTATCATATTTCTTGAATGTTGGTAAACATCTTTAGATTGAAGTTTCTCTCAGC 1300
Db 122762 TACCTTCTGTGTTTCTTAGGCGTGTGTCCTGTTCTGTTATGTTTCTGTTATT 122703
QY 1301 CTCTCTTT--AGTCTGCAATTTGAGATAGATATTTTACATCTGATTTATCTTAGAAT 1358
Db 122702 ATCTTTGAAGGCTGGATTTCTGAGAGATTAATGCGTAATTTGGTTTGTCTGGAAT 122643
QY 1359 GTCTTTCTTTCTCCAACTATTGTGACAGAAAGTTTTTCTAAGTGCAGTAGTCTGGCCTGA 1418
Db 122642 ACTTTGGTTTCTCCCTCTATGATAAATGAGAGTTTGGCTGGGTATAGTAGCCTGGGCTGC 122583
QY 1419 CATCTGTAGTCTCTGAGTCTGTAGCACATCTGTGAGGCGCTTCTTACATTTTGAAT 1478
Db 122582 AGTTTGTCTCTTAGTGTCTGTATAACATCTGTCCAGGCTCTCTGCGCTTTTCATAGTC 122523
QY 1479 TCTATTGGAAAAGTCAGGTGTAATTTCTAATACATCTGCTCTTTTATATGTTAATTTGCTCT 1538
Db 122522 TCTGCT-GAAAACTGCTGTAATTTCTGATAGGCTTGCCTTTATATGTTACTT-GACCTT 122465
QY 1539 TTTCCCTGCACTTTTAAATCTTTCTTTGTTCTATACCTTTTAGTGAATTTGATTA 1598
Db 122464 TTTCCCTACTGCTTTTATGATTTCTATCTTTTATTAGTCAATTTGATGTTCTGATTA 122405
QY 1599 TGCAGTGTGGGAGTTTCTTTTCCGGTCCCAATCTATTTGGTGTGTTTGTATGCTTCTTTGTA 1658
Db 122404 TGTGTGGGAGGAATTTCTTTTGTGTCAGTCTATTTGGAGTCTGTGAGCTTCTTTGTA 122345
QY 1659 CTTGATAGGCACTCTTTTCTCAAGGTTAGAAATTTTCTTTTGTGTTTCTTGA 1718
Db 122344 TGTTCATATGATCTCATCTTTAGATTGGGAAGTTTCTTCAATAATTTTGTGAAGA 122285

Db 21817 ATCCCTTTGTTTGGGATAAAATGTTCTGTAGATATCTGTTACATCCATTTTTCATTAAC 21876
QY 589 ATCAGTTAGTCCACATTTCTCTGTTTCGTTTTCGTTTGGTGGAGTACCTACTGCTGGAG 648
Db 21877 TTATGTTTGTATCAATATGTCCTGTTTGGTGGAGTATGTTTCTATTTCCATGATTTGCTCAATGATG 21935
QY 649 AGAATGGGATTTGAAGTAGCCCACTATCTGTGTGAGGT - CAATATGTAATTTAGCT 707
Db 21936 AAAGTGGTGTGTTGAAGTCCCCCACTATATTTGTGTGAGGTGCAATGTGTCTTTGAGCT 21995
QY 708 GTAGCTGTGCTGTTTATGAACTGGGTGACATGTTGTTGGTGCATAGACATTAAGAA 767
Db 21996 TTACTAAAGTGTCTTTAATGAATGTGCTGCTGTTGTTATTTGGAGCATAGATATTCAGAA 22055
QY 768 TTGCAATGTCTCTTTGGTGGATTTTCCCTTTGATGCTATGTAGTATTTCTTCCCAATCTCA 827
Db 22056 TTGAGAGTTCCTCTTTGGAGGATTTTATCTT - TGAGTATGAAGTCCCTCTCTTGTCTT 22113
QY 828 TCTGCTTGTGTTTGGGTTTAAAGTCTATTTAG - TCAGATATTAATAAGTCTATCGGCTTG 886
Db 22114 TTGTATGACCTTTGGGTTGGAGTGCATTTTATTCGATATTAAGAATGGCTACTCCAGCTTG 22173
QY 887 CTTCTTAGGGCAATTTGCTTAGAATA - TCTTTTCCATCTTTTACTCTTAAGGTGATGCT 945
Db 22174 TTTCTTCAGCAATTTGCTTGCAGAAATGTTTTCAGGCTTTCACTCTGAGGTAGTGTCT 22233
QY 946 ATCCAT - - - GGTAAGTGTCTTTTGTGATCAGCAGTAGATGGAATCTTGTCTTTCATAT 1002
Db 22234 GTCTTTATCCCTGAGATGGGTTTCTGTAAGCAGCAGAAATGTTGTGCTCTTTGTGTAG 22293
QY 1003 CCATTTCTGTACCCAGTATCTTTTCTAGAGAAATTAAGATCAITGAGTCAATGATGTTG 1062
Db 22294 CAGTTTGTGTAGTCTATGCTCTTTTATTTTGGGAGTTGAG - - - - - TCCATTTGATGTTA 22345
QY 1063 AGAATATCAATGAGCAGTGTGTTGGATCTTGTATTTCTGACATCTGTGAAGTGTGTG 1122
Db 22346 AGAGATATTAAGGAAAGTAAATGTTGCTTCTATATTTT - - - TTGTTAAAGTTGCA 22402
QY 1123 GTGCTGT 1182
Db 22403 TTTCTGTTCTGTGGCTGTCTTTTGTAGTTG - - - - - 22435
QY 1183 GTCTGTGTGTGTGTGTGTGTCTCTCCCTCTTTTGTATTTTGGCCTGGAATTTT 1242
Db 22436 - - - - - TTGAGGATTA 22446
QY 1243 ATTATTCATATTTCTGAAATGGGTAACATCTTTAGATTAAGTATTTTCTCTAGCCT 1302
Db 22447 CTTTCTTACATTTCTAGGATGTGTTTCTGCTCTGTTATTTGTTTCTGTTATAT 22506
QY 1303 TCTTT - - - AGGCTGATTTGAAGTAGATATTTCTTACATCTGATTTTATCTTGAATGT 1360
Db 22507 CTTTGAAGGACTGGAATTCGTGGAAGATAATGTTGTAATTTGTTTTCGTCGGAATAC 22566
QY 1361 CTTTCTTTCTCCAACTATTTGACAGAAAGTTTTTCTAAGTGCAGTAGTCTGGCCTGACA 1420
Db 22567 TTTGGTTTCTCCATCTATGTTAAATTTGAAAGTTGGCTGGGTATAGTAGCCTGGCTGGCA 22626
QY 1421 TCTGTAGTCTCTTGGAGTCTGTAGACATCTGTGAGGCTCTTACATTTTGAATTC 1480
Db 22627 TTTGTGTTCTCTTAGTGTCTGTATAACATCTGTCCAGATCTTTCTGCTTTTCATAGTCTC 22686
QY 1481 TATTGGAAGTCAAGTGTATTTCTTAATACATCTGCCCTTTATATCTTAATGTTCTTTT 1540
Db 22687 TGGT - GAAAGTCTTGTGTAATCTGTATAGGCTGCTTTATATGTTACTT - GACCTTTT 22744
QY 1541 TCCCTTGCATCTTTAAATATCTTTCTTTGTTCTATCTTTTGTAGTGAATTTATATG 1600
Db 22745 CCCCTTACTGTGTGAATATATATCTTTTATTTAGTGCATTTGTTGTCTGAATTTATG 22804
QY 1601 CACTGTGGGAGTTTCTTTTCCGGTCCCAATCTATTTTGGTGTGTTTGTATGTTCTTGTACC 1660

Db 22805 TGTCGGAGGAATTTCTCTTCTGGTTAGTCTATTTGGAGTTCTCTAGGCTTCTGTATG 22864
QY 1661 TTGATAGGCATCTCTTCTCAAGGTTAGGAAATTTTCTTTTGGTTTCTTGAATAA 1720
Db 22865 TTATGGGCATCTCTCTCTTTAGGTTGGGAAGTTTCTTCTATATTTTGTGACGATA 22924
QY 1721 TTTTCCCTCTTTTGGACCTG - - - - - CCTTCTCCCTTCTCTATTTCCCTTT 1766
Db 22925 TTTGCTGGCCCTTTAAGTTGAAATCTTCAATCTCATCTACTCCCATATCCGTAGGTTT 22984
QY 1767 GGTTTTGTATAGTGTCTCTGCTTCCCTGATGTTTATGCTCGGATTTATTTAGACTTA 1826
Db 22985 GGTCTTCTGTGTGCTCTGGATTTCTCGGATGTTTAAAGTTAGGATCTTTTGGCATTT 23044
QY 1827 ACATTTTCTTTGACCAAGGTATCCATTTCTTCTATCTTGTCTTCTGCTGAGATTTCTC 1886
Db 23045 GCATTTGCTTTGATTTGTTGCGGATGTTCTCTATGGAATCTCTGTACTGAGATTTCTC 23104
QY 1887 TCTTCTATCTCTTGTATCTGTCAGTGGGCTGCTCTGAGGTTCCCTG - TTGGGTTCTT 1945
Db 23105 TCTTCCATTTCTTTGATTTCTGTTGTGATGCTCACATCTATGTTCCAGATTTCTTTCT 23164
QY 1946 AATTTTTTCAATTTCCAGATTTCTCTTCAAGTTTGGGTTTTGTTTATTAATCTATTTCCACT 2005
Db 23165 AGGTTTCTATCTCCAGGTTGCTTCCCTCACCTTTGGGTTTTCTTTATTTGTGCTCACTTCTCT 23224
QY 2006 TTCAGTCTCTGAATGTTTACTCAATTTTCTCCAGGATTTTAC - - - - - TT 2052
Db 23225 TTAGTCTAGTATGTTTGTTCATTTTCCATCACTGATGTTGGGATGTTTCTCTGTT 23284
QY 2053 TTCATAGGTTTCTTAAATGGAATTTATCAATTTCTCTTCAAGGACCTTTTATGAATTCAT 2112
Db 23285 TTTCTTAAAGGACTTGTAACTCTTTAGCAGTGTCTCTCTGATTTCTTTAAGTGAGTTAT 23344
QY 2113 AAAATGATGTTAAGTCTCT - TGCCTTGTGCTTCAAGTATGTTGCAATCTCTCAGGCTTAT 2171
Db 23345 AAAGTCTTCTTGTATGCTCTCTACCATCATCATGATATGCCCTTTAAATCCAGTCTAG 23404
QY 2172 TGTAAATAGGTTTGTAG 2187
Db 23405 CTTTGGGTTTTTGG 23420

RESULT 18
US-10-003-806-10
; Sequence 10, Application US/10003806
; Publication No. US2002011929A1
; GENERAL INFORMATION:
; APPLICANT: Bishop, Colin E.
; APPLICANT: Agoulnik, Alexander I.
; APPLICANT: Zhu, Qichao
; TITLE OF INVENTION: CAN 1 AND ITS ROLE IN MAMMALIAN INFERTILITY
; FILE REFERENCE: P02066U1/10024824
; CURRENT APPLICATION NUMBER: US/10/003,806
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: 60/245,872
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 173808
; TYPE: DNA
; ORGANISM: Mouse
US-10-003-806-10

Query Match 25.4%; Score 578.299; DB 1; Length 173808;
Best Local Similarity 63.2%; Pred No 1.5;
Matches 1147; Conservative 0; Mismatches 547; Indels 121; Gaps 18;
QY 292 TCTATTTCTTGAATTTCTATCTTTGGGCTCAATTTTAACTCAGTAGTGAGTTGTTGTTTCC 351
Db 159106 TTTCTTTCTTTATCTTCTTCCCTTCAAGGATATCATTCAGAGAGAGTGTGTTTCACTTTCC 159165

OTHER INFORMATION: n
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (3503)..(3503)
 OTHER INFORMATION: n
 US-10-388-934-223

Query Match 25.4%; Score 578.2; DB 1; Length 8048;
 Best Local Similarity 62.0%; Pred. No. 5.8;
 Matches 1160; Conservative 0; Mismatches 574; Indels 136; Gaps 17;

292 TCTATTTCTGATTTCTATCTGGCTCATTTTAACTCAGTAGAGTGTGTTGGTTTCC 351
 3329 TTTCTTTCTTTATTTCTTCTGACCAAGTTATCATTTAGTAGAGCATTTCAATTTCC 3270
 352 ATAGTTTGTAAAGTTTCTGTTGTTTCTGTTGTTGTTGTTGTTATCTAGATTTAAAGCTGT 411
 3269 AGGTATATGTGGCAATTC-----TTCCCTTATTTGTTATTTGAACACAGTTTAAAGCCGT 3216
 412 GGTGTCAGATAGACATAGATATTTTCAATTTCTTTTATCTGTCGAGACTTTCCTT 471
 3215 GGTGTCGATAGACGATAGGATATTTCTATCTTCTGATCTGTTGAGGCCGTTT 3156
 472 TGTTTTGAATATGATTTCAATTTTGGAG--AGTTTCATAGGTTGCTGACAGAGAGTAC 529
 3155 TTTGACCAATATATGTTCAATTTTGGAGAAAGTCCCATGAGGAGCTGAGAGAAAGGTAT 3096
 530 AGTCTTTGTTTGGTGAATAGTCTGTAAA--TATCTCTAGTCCACTTTGTTTATGA 587
 3095 ATCCTTTGTCTTAGGATAGATGTTCTAAATATCGTTAAGTCCATTTGGCTCATGA 3036
 588 CATCAGTTAGTCCAGCATTTCTGTTTCTGTTTCTGTTTGTGAGATGACCTAACTGTTGGA 647
 3035 CTCTCTTATGCTGTCGACATCACTGTTT-AAATTTCTGTTTCCATGATCTGTCCATTTGAT 2977
 648 GAGATGGGTATTTGAAGTACCCACTATCTGTGTGTGAGGT-CAATATGTTGATTTTACG 706
 2976 GAGAGTGTGTGTTGAATCTCCCACTATTTATTTGTGTGAGGTGCAATGTGTGCTTTGAGC 2917
 707 TGTAGCTGTGCTGTTTATGAACTTGGGTGACATGTTGTTGTTGTCATAGACATTAAGA 766
 2916 TTTAGTAAGGTTCTTTTACGTATGATGTCCTGTTGTTTATTTGGGCATAGATATTTAGG 2857
 767 ATTCAGATGCTCTTGGTGA-TTTTCTTTGATGCTATGTAGTATTTCTTCCCAATCT 825
 2856 ATTGAGAGTTTCACTGTTGTTGTTTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2797
 826 CATCTGCTTATTTGGGTTTAACTCTA-TTATGTCAGATATTAATAAGTCTATCGGCT 884
 2796 TTTTGTATGACTTTTATGTTGAAATTTGATTTTATTTGATATTTAGATGTTCTTCCAGCT 2737
 885 TGCTTCTTATGGCCATTTGCTTAG-ATAATCTTTTCCATCTTTTACTCTTAAGTGTATGT 943
 2736 TGCTTCTTCTGACCATTTTGTGTTGTTGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2677
 944 CTATCCATG---GTAGGTTGTTTGTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1000
 2676 CTGCTTTGCTCTGAGGTGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2617
 1001 ATCCATTTCTGTTACCCAGTATCTTTTCTGAGAGAAATTAAGATCAATTTGATGATGTT 1060
 2616 ATCCAGTTTGTAACTATGACTTTTATTTGGGAGTTGAG-----GCCATTTGATGT 2565
 1061 TGAGAAATTAATCAATGAGCAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1120
 2564 TGAGAGATATTAAGGAATAGTATGTTGTTTCCGTTTATTTCAATTTTGGATTTGAGG 2505
 1121 GT 1180
 2504 TTAGTTTGTGTG----- 2492
 1181 GT 1240

2491 -----CTTTCATTTCTTTGTTTGTGTTGCCAAGACGAT 2459
 1241 TTATTTATTCATATTTTCTGAAATGTGGTAACATCTTTAGATTTGAAGTTTCTCTCCTAGC 1300
 2458 TAGTTTCTGCTTCTTTCTAGGATAGCTTGGCTCCTTATGTTGGCTTTTACCATTTAT 2399
 1301 CTTCCT--TAGGTCTGCATTTGAAGATAGATATTTCTTACATCTCTGATTTTATCTTAGAAT 1358
 2398 ATCCCTTTGTAGTGTGATTTTGTGAAAGATATTTGTAAATTTTGGTTTGTGTCATGGAAT 2339
 1359 GTCCTTTCTTCTCCAACTATTGTGACAGAAAGTTTCTTAAAGTGCAGTAGTCTGSCCTGA 1418
 2338 GTCCTGTTTCTCCATCAATTTAGAGTTTGTGGAATACAGTAACCTGGGCTGG 2279
 1419 CATCTGAGTCTCTGGAGTCTGTAGCACATCTGTGCGAGGCTTCTTACATTTTGAAT 1478
 2278 CATTTGTGTTCTTAGGCTCTGTATACATCAGTCCAGGATCTTCTGGCCTTCATAGTT 2219
 1479 TCTATTTGGAAGTGCAGGTGTAATTTCTTAATACATCTGCTTTTATATGTATTTGTTCTTT 1538
 2218 TCT-GGCGAGAAAGTCTGTTGTAATTTCTGATAGGCTCTCCCTTTATATGTACTT-GACCTT 2161
 1539 TTTCCCTTGCATCTTTTAAATATTTCTTCTTCTTCTATCTATCTTTTATGATTTGATTATTA 1598
 2160 TTTCCCTTACTGTTTAAATATCTTTCTTATTTTGTGCGTTTGGTGTGTTGACAAATTA 2101
 1599 TGCACGTGTTGGGAGTTTCTTTTCCGTTCCAAATCTATTTTGGTGTGTTTGTATGCTTCTTGA 1658
 2100 TGTGACGGGAGGTGTTCTTTTCTGTTCCAAATCTATTTGGAGTTCTGTAGGCTTCTTGTG 2041
 1659 CTTGTAGTAGCATCTTTTCTCAAGTTAGGAAATTTTCTTTTGTGTTTGTGTTTCTTGA 1718
 2040 TGTCTATGGGTATCTTTTATTTAGTTAGGAAAGTTTCTTCTATGTTTGTGTTGAAGA 1981
 1719 TATTTTCCCTGCTTTTGTGACCTG-----CCTTCTTCCCTTCTCTTATTTCT 1764
 1980 TATTTTACTGCTTCTTGTAGCTGGAGTCTTCACTCTCTTCTATACCTATTTATCTTAGGT 1921
 1765 TTGTTTGTGATGTTGCTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 1824
 1920 TTGATCTTCTCATTTGAGTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 1861
 1825 TAACATTTTCTTTGACCAAGGTATCAATTTCTTCTATCTTCTTCTTCTTCTTCTTCTTCTTCT 1884
 1860 TTACATTTCTTTGACGTGAGTCAATGATTTCTATGGAATCTTCTGCTCTGCTGATTC 1801
 1885 TCTCTTCTATCTTCTTGTATTTCTGTCAGTGGCTTGTCTGAGGTTCTGCTGTTGGGTTCT 1944
 1800 TCTCTTCTCATCTTGTATTTCTGTTGTTGAAAGCTTTTATCTACAGCTCTTCTGCTTCTTCT 1741
 1945 TAAATTTTCT-ATTTTCCAGATTTCTTCTTCTGTTTGGTTTGTGTTTATTAATTTCTATTTCCA 2003
 1740 TTTGGTTTCTATATCCAGGTTGTTTCCATGTTGTTTCTTTTGTGTTGTTGTTGTTGTTGTTGTT 1681
 2004 CTTTCAGGTCCTGAAATGTTTACTTCAATTTCTCTCCAGTATTTTACATTTTCTCATAGGTTT 2063
 1680 TTTTAAATTTCTTCAACTGTTT-----GATTTGTTTCTTCTGGAATTC 1638
 2064 CTTTATGATTTTATTTCAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2123
 1637 CTTTCAGGATTTTGTGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1578
 2124 TAAAGTCTTCT 2133
 1577 GGAATTTCTTT 1568

RESULT 20
 US-10-388-934-224/c
 ; Sequence 224, Application US/10388934
 ; Publication No. US20040005547A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Boess, Franziska

QY 2004 CTTTCAGTCTCGAATGTTTACTCATTTTCCTCCCAAGTATTACATTTTCATAGGTTT 2063
 Db 1680 TTTTAAATTCCTCAACGTTT-----GATTGTGTTTCTCGAATTC 1638

QY 2064 CTTTAAATGATTTATTCATTTCCCTTCACAGGACCTTTATGAATTCATATAAATGATGT 2123
 Db 1637 CTTTCAGGATTTTGTGTCCTCTCTATGCGCTTCTACTTGTTTATTATGTTTCTT 1578

QY 2124 TAAGTCTCTT 2133
 Db 1577 GGAATCTTT 1568

RESULT 21
 US-10-388-934-682/c
 ; Sequence 682, Application US/10388934
 ; Publication No. US2004000547A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Boes, Franziska
 ; APPLICANT: Super-Dick, Laura
 ; APPLICANT: Wolf, Detlef
 ; TITLE OF INVENTION: BIOMARKERS AND EXPRESSION PROFILES FOR TOXICOLOGY
 ; FILE REFERENCE: 21199
 ; CURRENT APPLICATION NUMBER: US/10/388,934
 ; CURRENT FILING DATE: 2003-03-14
 ; PRIOR APPLICATION NUMBER: 02005336.9
 ; PRIOR FILING DATE: 2002-03-14
 ; PRIOR APPLICATION NUMBER: 02015657.6
 ; PRIOR FILING DATE: 2002-07-17
 ; NUMBER OF SEQ ID NOS: 862
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 682
 ; LENGTH: 8048
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus (No. US2004000547A1way rat)
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (2005)..(2005)
 ; OTHER INFORMATION: n
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (3503)..(3503)
 ; OTHER INFORMATION: n
 US-10-388-934-682

Query Match 25.4%; Score 578.2; DB 1; Length 8048;
 Best Local Similarity 62.0%; Pred. No. 5.8; 574; Indels 136; Gaps 17;
 Matches 1160; Conservative 0; Mismatches

QY 292 TCTATTTCTGATTTCTATCTGGCTCATTTTAACTCAGPAGTGGTGTGTTGTTTC 351
 Db 3329 TTTCTTTCTTTATTTCTTCTGACAGGTTATCATTCAGTAGAGCATTTGTTCAATTTCC 3270

QY 352 ATAAGTTTGTAAATTTCTGTTGTTTCTGTTGTTGTTGTTGTTTATCTAGATTAAAGCTGT 411
 Db 3269 AGGTATATGTGGCAATTC-----TTCCCTTATTTGTTATTGAAGCAGGTTTATAGGCCGT 3216

QY 412 GGTGTCAGATGAGACATAGATATTATTTCAATGTCCTTTATCTCTGTCAGACTTGCCTT 471
 Db 3215 GGTGTCGATAGACGATGAGATTATTTCTATCTTCTGATCTGTTGAGGCCGCTT 3156

QY 472 TGTGTTGAAATGATGATTTCAATTTTGGAG--AGTTTCATAGGGTGGTGCACAAGAGTAC 529
 Db 3155 TTTGACCAATTATATGTTCAATTTTGGAGAAAGTCCCATGAGGAGCTGAGAAGAGGTAT 3096

QY 530 AGTCTTGTGTTTGGTGAATAGTCTGTAAA--TATCTCTAGTCCACTTGGTTTATGA 587
 Db 3095 ATCCTTTGCTTTAGGATAGATGTTCTAAANAATATCGGTTAAGTCCATTTGGCTCATGA 3036

QY 588 CATCAGTATGATCCAGCATTTCTCTGTTTCGTTTTTTTGTGAGATGACCACTTAACGTGTGGA 647
 Db 3035 CTTCTCTATGCTGTCGACATCACTGTTT-AAATTTCTGTTTCCATGATCTCTCCATTGAT 2977

QY 648 GAGATGGGGTATTGAAGTAGCCCACTATCTCTGTGTGAGGT-CAATATATGATTTTATGC 706
 Db 2976 GAGAGTGGTGTGTGTAATCTCCCACTATTATTTGTGTGAGGTGCAATGTGTCTTTGAGC 2917

QY 707 TGTAGCTGTGCTTGTGTTTATGAACCTTGGGTGACATTTGTTTGGTGCATAGACATTAAGA 766
 Db 2916 TTTAGTAAAGTTCCTTTACGTATGTAGTGTCCCTTGTTATTTGGGCGCATAGATTTAGG 2857

QY 767 ATTGCAATGCTCTCTGTGTGGA-TTTTCTTTGATGCCCTATGATGATTTCTTTCCAAATCT 825
 Db 2856 ATTGAGAGTTCATCTGTGTGGATTTTCTTTGATGAATATGAAGTGTCTTCTCTATCT 2797

QY 826 CATCTGCTTAGTTTGGGTTTAAAGTCTA-TTAGTCAGATATTAAATGACGTATCGGCT 884
 Db 2796 TTTTGTATGACCTTTAGTTTGGAAATTTGATTTTATTGATATTAGAAATGGCTACTCCAGCT 2737

QY 885 TGTCTTTAGGGCCATTTGCTTTAG-AAATATCTTTTCCATCTCTTTTACTCTAAGGTGATGT 943
 Db 2736 TGTCTTCTCTGACCAATTTGCTTTGGAAATTTTGTTTTCCAGCCTTTTCACTCTGAGGTAGTGT 2677

QY 944 CTATCCATG---GTAGGTTGTCTTTTGGATGACAGCAGTAGGATGGATCTTGTTTTCAT 1000
 Db 2676 CTGTCTTTGTCTCTGAGGTGTCTTCTGTAGGCGACAGAAATGCAGGGTCTCTGTGTGTGT 2617

QY 1001 ATCCCATTTGTTACCAGTATCTTTTCTAGAGAAATTAAGATCATTCAGTCAATTGATGT 1060
 Db 2616 ATCCAGTTTGTAACTATGACTTTTATTTGGGAGTTGAG-----GCCATTTGATGT 2565

QY 1061 TGAGAAATTAATCAATGACGAGTGTGTTGTTGATCTTTGTTATCTGTCACCTGTGAAAGTGTGT 1120
 Db 2564 TGAGAGATATTAAAGGAATAGTGAATTTGCTTTCCGCTTATATTTCAATTTTGGATGTGAGG 2505

QY 1121 GT 1180
 Db 2504 TTAGTTTGTGTGTG----- 2492

QY 1181 GTGTCTGTGTCTGTGTGTGTGTGTCTCTCCCTCTTTTGTATTTTGGCTTGGAAATPAT 1240
 Db 2491 -----CTTTCAATCTCTTTTGTGTTTGTGCCAAGCAT 2459

QY 1241 TTATTTATTCATATTTCTTGAATGGGTACATCTTTAGATTGAAGTTTTCCTTCTTAC 1300
 Db 2458 TAGTTTCTTGTCTCTCTAGGATAGCTTGGCTCTTATGTGCGCTTTTACCATTATT 2399

QY 1301 CTTCTTT--TAGGTCTGCATTTGAAGATAGATATCTTTTACATCTGATTTTATCTTAGAAT 1358
 Db 2398 ATCTTTTGTAGTGTGGATTTGTGGAAGATATTCTGTAAATTTGTTTGTCTATGGAAT 2339

QY 1359 GTCTTTCTTTCTCCAACTATTGTGACAGAAAGTTTCTTAAGTGCAGTAGTGTGCCCTGA 1418
 Db 2338 GTCTTGGTTTCTCCATCAATGTTAAATTTGAGAGTGTGCTGGATACAGTAACCTGGGCTGG 2279

QY 1419 CATCTGATCTCTTGGAGTCTGTAGCACATCTGTGAGGGCTTCTTACATTTTGTAGTT 1478
 Db 2278 CATTTGTGTTCTTCTAGGTCGTATACATCAGTCCAGGATCTTCTGGCTTCTATAGTT 2219

QY 1479 TCTATTGGAAGAGTCAGGTGTAATTTCTAATACATCTGCTTTTATATGTTAATTTGTCPTT 1538
 Db 2218 TCT-GGCGAGAAAGTCTGTGTGATTCTGATAGGTCTCCCTTTATATGTTACTT-GACCTT 2161

QY 1539 TTTTCCCTTGCATCTTTTAAATTTCTTTCTTCTATCTATCTTCTTCTTCTTCTTCTTCTT 1598
 Db 2160 TTTCCCTTCTCTGTTTAAATTTCTTCTTATTTTGTGCGTTTGGTGTGTTTGACAATA 2101

QY 1599 TGCACGTGTGGGAGTCTCTTTTCCGCTCCAATCTATTTTGGTGTGTTTGTATGCTTCTTCTGA 1658
 Db 2100 TGTGAGGAGGAGTGTCTTTTCTTCTGTTCCAATCTATTTGGAGTCTGTAGGCTTCTTGTG 2041

QY 1659 CCTTGATAGGATCTCTTCTCAAGTTAGGAATTTTCTTTTGTGTTTCTTCTTCTTCTTCTT 1718
 Db 2040 TGTCTATGGGTATCTCTTTTATAGTTAGGAAGTTCCTTCTATGTTTGTGTTGTTGTTGTTG 1981

QY 1719 TATTTTCTCTGCTTTTGGACCTG-----CCTTCTTCCCTTCTCTTATTCCT 1764

Db 2278 CATTGTGTTCTTAGGGTCTGTATAACATCAGTCCAGGATCTCTGGCCTTCATAGTT 2219
QY 1479 TCTATTGGAAAGTCAGGTGAATCTAAATACATCTGCCCTTATATGTTAAATGGTCTTT 1538
Db 2218 TCT-GGGAGAGAGTCTGGTGATCTGTAGAGTCTGCCCTTATGATGTTACTT-GACCTT 2161
QY 1539 TTTCCCTTGCATCTTTTAAATATCTTTCTTTGTTCTATACATTTTATGATGATTA 1598
Db 2160 TTTCCCTTACTGTTTAAATATCTTTCTTTTATTTGCGGTTTGGTGTGTTGACAAATTA 2101
QY 1599 TGCAGTGGGAGGTTCTTTTCCGGTCCCAATCTATTGTTGTTGTTGATCTCTTGTA 1658
Db 2100 TGTGACGGGAGGTGTTCTTTTCTGGTCCCAATCTATTGAGTCTCTAGGCTCTTTG 2041
QY 1659 CTTTGATAGGATCTCTTTCTCAAGGTTTAGGAAATTTTCTTTTGGTTTTCTTGAAAA 1718
Db 2040 TGTCTATGGGTATCTTTTATTTAGGTTAGGAAAGTTTCTTCTATGTTTGTGGAAGA 1981
QY 1719 TATTTTCCCTGCTTTGACCTG-----CCTTCTCCCTTCTCTATTCCT 1764
Db 1980 TATTTACTGTCCTTTGAGCTGGAGTCTTCACTCTCTCTATACCTATTATCCTTAGGT 1921
QY 1765 TTGGTTTTGTCATAGTGTCTCTGGCTTCTCGATGTTTATGCTCGGATTTATTTAGACT 1824
Db 1920 TTGATCTCTCATGAGTCTGGATTTCTCTGTATGTTTGGACAGTAGCTTTTCCGCT 1861
QY 1825 TAACATTTTCTTTGACCAAGGATTCATTTCTTCTATCTGTTCTTCACTGCTGAGATTC 1884
Db 1860 TTCAATATCTTTGACAGTCTGAGTCAATGATTTCTATGGAATCTCTGCTCTGAGATTC 1801
QY 1885 TCTCTTCTATCTGTGTTATCTGTCAGTGAGGCTTGTCTGAGGTTCCCTGTTGGGTTCT 1944
Db 1800 TCTCTTCACTCTCTGTGTTATCTGTTGGTGAAGCTTTTATCTAGAGTCTCTGCTCTCT 1741
QY 1945 TAAATTTTTC-ATTTCCAGATTTCTTCCAGTCTGGGTTTTGTTTATTAATCTATTTCCA 2003
Db 1740 TTTGGTTTTCTATATCCAGGGTGTGTTCCATGTGTCTTTTTCATTTGCTCTATTTCCA 1681
QY 2004 CTTTCAGTCTCGAAAGTTTTACTCATTTTCTCCAGATTTTACATTTTTCATAGTTT 2063
Db 1680 TTTTAAATCTCTCACTGTTT-----GATTTGTTTCTCTGGAATTC 1638
QY 2064 CTTTAAATGATTTATTTCTCTCTCAAGAGCTTTTATGAATTCATAAATGTATGT 2123
Db 1637 CTTTCAGGATTTTGTGTCCTCTCTATGGGCTTCTACTTGTGTTATTAATGTTTCT 1578
QY 2124 TAAGTCTTT 2133
Db 1577 GGAATCTTT 1568

RESULT 23

US-10-388-934-785/c
; Sequence 785, Application US/10388934
; Publication No. US20040005547A1
; GENERAL INFORMATION:
; APPLICANT: Boess, Franziska
; APPLICANT: Suter-Dick, Laura
; APPLICANT: Wolf, Detlef
; TITLE OF INVENTION: BIOMARKERS AND EXPRESSION PROFILES FOR TOXICOLOGY
; FILE REFERENCE: 21199
; CURRENT APPLICATION NUMBER: US/10/388,934
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: 02005336.9
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 02015657.5
; PRIOR FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 862
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 785
; LENGTH: 8048
; TYPE: DNA
; ORGANISM: Rattus norvegicus (No. US20040005547A1way rat)

FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2005)..(2005)
; OTHER INFORMATION: n
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3503)..(3503)
; OTHER INFORMATION: n
; US-10-388-934-785

Query Match 25.4%; Score 578.2; DB 1; Length 8048;
Best Local Similarity 62.0%; Pred. No. 5.8;
Matches 1160; Conservative 0; Mismatches 574; Indels 136; Gaps 17;

QY 292 TCTATTCTTGATTTCTATCTCTGGCTCAATTTTAACTCAGTAGTCAGTTGTTGTTTCC 351
Db 3329 TTTCTTTCTTTATTTCTTCTTGACAGGTTATCATTCAGTAGAGCATTTGTTCATTTCC 3270
QY 352 ATAAGTTTGTAAAGTTTCTGTTGTTTCTGTTGTTTCTGTTGTTTCTATCTAGATTAAAGCTGT 411
Db 3269 AGGTATATGTTGGCAATC-----TTCCCTTATTTGTTATGAAGACCACTTTTAGCCGT 3216
QY 412 GGTGTCAGATAGGACATAGAGTATTTTCAATTTGCTTTTATCTGTCGAGACTTGTGTT 471
Db 3215 GGTGGTCCGATAGCACGATGGGATTTATTTCTATCTTTCTGTACCTGTGAGGCCCGTT 3156
QY 472 TGTTTTGAATATGTTATTCATTTTGGAG--AGTTTTCATAGGGTCTGCACAAGAGGTAC 529
Db 3155 TTGACCAATATATGTTCAATTTTGGAGAAAGTCCCATGAGAGCTGAGAGAGAGGTAT 3096
QY 530 AGTCTTTGTTGTTTGGTGAATATGTTCTGTAA--TATCTAGTCTCCACTTGGTTTATGA 587
Db 3095 ATCTTTTCTTTAGGATAGATGTTCTAAATAATATCCGTTAAGTCCATTTGGCTCATGA 3036
QY 588 CATCAGTTAGTCCAGCAATTTCTGTTTCTGTTTCTGTTTCTGATGAGTCACTTCTGTTGA 647
Db 3035 CTCTCTTATGCTGTGACATCACTGTTT-AAATTTCTGTTTCTCCATGATCTGTCCATGAT 2977
QY 648 GAGATGGGTTATGAAGTAGCCCACTATCTGTGTGAGGT-CAATATGATGATTTTAC 706
Db 2976 GAGAGTGGTGTGTTGAATCTCCCACTATTATGTTGAGGTGCAATGTGCTTTGAGC 2917
QY 707 TGTAGCTGTCTGTTTATGAACCTTGGGTGACATTTGTTTGGTGTGATAGACATTAAGA 766
Db 2916 TTTAGTAAGTTTCTTTTACGTATGTAGTGGCTTGTATTTGGGGCATAGATATTTAGG 2857
QY 767 ATTCGAATGCTCTTGTGGTGA-TTTTCTTTGATGCTCTATGTATTTCTTCCCAATCT 825
Db 2856 ATTGAGAGTTCACTTTGGTGGATTTTCTTTGATGAATGAAGTGTCTTCTTCTTACT 2797
QY 826 CATCTGCTTAGTTTGGGTTTAAAGTCTA-TTAGTCAGATATTTAAATGACGTATCGGCT 884
Db 2796 TTTTTCATGACTTTTAGTTGGAATTTGATTTATTTAGATTTAGAAATGGCTACTCCAGCT 2737
QY 885 TGCCTTTAGGGCCATTTGCTTAG-AAATATCTTTTCCATCTTTTACTTCTTAAGGTGATGT 943
Db 2736 TGCCTTCTCTGACCAATTTGCTTTGGAAATTTGTTTTCAGCCCTTTTCTCTGAGGTAGTGT 2677
QY 944 CTATCCATG---GTAGGTTGCTTTTTTGGATGTCAGCAGTAGGATGATGATTTTTCAT 1000
Db 2676 CTGCTTTGCTCTGAGGTGTGTTTCTCTAGGACAGATGACAGGGTCTCTGTTGTGT 2617
QY 1001 ATCCATCTGTACCAGCATCTTTTCTTAGAGAAATTAAGATCATTTAGTCAATGATGT 1060
Db 2616 ATCCAGTTTGTAAATCATGACTTTTATTTGGGAGTTGAG-----GCCATGATGT 2565
QY 1061 TGAGAAATTAATCAATGAGCAGTGTGTTGATTTCTTTTATCTTGCATTTGCACTTGTGAAGTGT 1120
Db 2564 TGAGAGATTAATAGGAATAGTATTTATCTTCCCGTTATATTCATATTTGATGTGAGG 2505
QY 1121 GT 1180
Db 2504 TTTATGTTTGTGTG----- 2492


```

QY 885 TGCTTCTTAGGCCATTTGCTTAG-AAATCTCTTTTCCATCCTTTTACTCTAAGGTGATGT 943
Db 2736 TGCCTCTTCTGACCAATTTGCTTGGAAATTTGTTTTCCAGCCTTTTCACTCTGAGGTAGTGT 2677
QY 944 CTATCCCATG---GTAGGTGTCTTTTTTGGATGAGCAGTAGGATGAGTCTGTTTTCAT 1000
Db 2676 CTGCTTTTGTCTCTGAGGTGTCTTCTGTAGGCAGCAGATGAGGCTCTCTGTGTGT 2617
QY 1001 ATCCATCTCTTACCAGTATCTTTTCTAGAGAAATTAAGATCATTTGAGTCATTGATGT 1060
Db 2616 ATCCAGTTTGTAACTATGACTTTTATTTGGGAGTTGAG-----GCCATTTGATGT 2565
QY 1061 TGAGAAATATCAATGAGCAGGTGTGTGGATTTCTTTTATCTTGCACCTTTGTGAAGTGTGT 1120
Db 2564 TGAGAGATATTAGGAATAGTATGATTTATGCTTCCGTTATATTCATATTTCGATGTGAGG 2505
QY 1121 CTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1180
Db 2504 TTAATGTTTGTGTG----- 2492
QY 1181 GTGTCTGTGTCTGTGTGTGTGTGTCTCTCCCTCTTTTGATTTTGGCTGGAATAT 1240
Db 2491 -----CTTTCATCTCTTTGTTTGTGGCAGACGAT 2459
QY 1241 TTAATTAATCATATTTCTTGAATGGGTAAACATCTTTTAGATTTGAAGTTTTCCTCAGC 1300
Db 2458 TAGTTTCTGTCTCTCTTCTAGGTATAGTTCCTCTTATGTTGGGCTTTACCAATTTAT 2399
QY 1301 CTTCTT--TAGGTCTGATTTTGAAGATAGATTTCTTTTACATCTGATTTTATCTTGAAT 1358
Db 2398 ATCCCTTTGTAGTGTGGAATTTGTGGAAGATATTTGTGAAATTTGGTTTGTCTATGAAT 2339
QY 1359 GTCTTTCTTTCTCCAACTATTGTGACAGAAAGTTTCTTAACTAGTCAGTAGTCTGGCTGA 1418
Db 2338 GTCTGTGTTTCTCCATCAATGTTAATTGAGAGTTTGTCTGATACAGATTAACCTGGCTGG 2279
QY 1419 CATCTGTAGTCTCTTGGAGTCTGTAGACACATCTGTGAGGGCCCTTCTACATTTTGATTT 1478
Db 2278 CATTTGTGTTCTTCTTAGGTCTGTATAACATCAGTCCAGGATCTCTCTGGCTTCTCATAGT 2219
QY 1479 TCTATTGGAAGAGTCAGGTGAATTTCTAATCATCTGCTCTTATATGTTAATTTGGTCTTT 1538
Db 2218 TCT-GCCGAGAGTCGTGTGATTTCTGATAGTCTCCCTTTATATGTTACTT-GACCTT 2161
QY 1539 TTTCCCTTGCATCTTTTAAATCTCTTTCTTTCTATATCTTTTAGTATGATTTGATTA 1598
Db 2160 TTTCCCTTACTGTTTTTAAATCTCTTTCTTTATTTGTGCGTTTGTGTTTGTGCAATTA 2101
QY 1599 TGCATCTGGGGAGTCTTTTCCGGTCCAACTATTTGTTGTTGTTTGTATGCTTCTGTGA 1658
Db 2100 TGTGACGGGAGGTCTCTTTCTGCTGCTCAATCTATTTTGGAGTTCTGTAGGCTTCTTTGTG 2041
QY 1659 CTTTCATAGGCATCTCTTTCTCAAGGTTAGGAATTTTCTTTTGTGTTTCTTGAAAA 1718
Db 2040 TGTCTATGGTATCTCTTTTATAGTTTAGGGAAGNTTCTCTATGCTTTTGTGTTGAAGA 1981
QY 1719 TATTTTCCCTGCTTTTGACCTG-----CCTTCTCCCTCTCTATTCCT 1764
Db 1980 TATTTTACTGGTCTTTGAGCTGGGAGTCTTCACTCTCTCTATACCTATTAATCTTAGGT 1921
QY 1765 TTGGTTTTTGTATAGTGTCTCTGGCTCTCGATGTTTTTATGCTCTGGATTTATTTTAGACT 1824
Db 1920 TTGATCTCTCATTCAGTCTCGAATTTCTGTATGTTTGGACAGTAGCTTTTTCGCT 1861
QY 1825 TAACTATTTCTTTGACCAAGGTATCCATTTCTTCTATCTTGTCTTCACTGCTGAGATTC 1884
Db 1860 TTACATTAATCTTTGACAGTTGAGTCAATGAATTTCTATGGAATCTTGTCTCTGAGATTC 1801
QY 1885 TCTCTTCTATCTCTTGTATCTGTCTGAGTGGCTTGTCTCTGAGTCTCTGCTGCTGCTGCT 1944
Db 1800 TCTCTTCCATCTTTGTATCTGTGTGGTGAAGCTTTTATCTACAGCTCTCTGTCTCTCT 1741
QY 1945 TAAATTTTTTCT-ATTTCCAGATTTCTCTCAGTTTGGGTTTGTATTAATTTCTATTTCCA 2003

```

```

Db 1740 TTTGGTTTCTATATATCCAGGGTTGTTTCCATGTGTCTTTTGTATGCTTCTATTTCCA 1681
QY 2004 CTTTCAGGTCTCGAATGTTTGTACTCATTTTCCCTCCAGTATTTACATTTTTCATAGGTTT 2063
Db 1680 TTTTAAATTCCTTCAACTGTTT-----GATTGTGTTTCTCTGGAATTC 1638
QY 2064 CTTTAAATGGAATTTATCATTTTCTCTTCAAGGACCTTTTATGAATTCATAAAATGTATGT 2123
Db 1637 CTTTCAGGATTTTGTGTCTCTCTCTATGGGCTTCTACTTGTGTTTATGTTTCTCT 1578
QY 2124 TAAGGTCTCTT 2133
Db 1577 GGAATTCCTT 1568

```

RESULT 25
 US-10-087-192-547/c
 ; Sequence 547, Application US/10087192
 ; Publication No. US20020182586A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Morris, David W.
 ; APPLICANT: Engelhard, Eric K.
 ; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
 ; TITLE OF INVENTION: CANCER
 ; FILE REFERENCE: 529452000122
 ; CURRENT APPLICATION NUMBER: US/10/087,192
 ; CURRENT FILING DATE: 2002-03-01
 ; PRIOR APPLICATION NUMBER: US 09/747,377
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: US 09/798,586
 ; PRIOR FILING DATE: 2001-03-02
 ; NUMBER OF SEQ ID NOS: 2059
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 547
 ; LENGTH: 116858
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(116858)
 ; OTHER INFORMATION: n = A,T,C or G
 US-10-087-192-547

Query Match 25.2%; Score 574.999; DB 1; Length 116858;
 Best Local Similarity 62.1%; Pred. No. 1.8; Indels 124; Gaps 18;
 Matches 1150; Conservative 0; Mismatches 555;

```

QY 292 TCTATTCTTGTATTTCTATCTTTGGCTCATTTTAACTCAGTAGTGAGTTGTTTCTCC 351
Db 37530 TTTCTTTCTTTATTTCTTCTTCCCTTGTATCAAGGTATCATTTAGTAGAGTTGTTTCTCC 37471
QY 352 ATAAGTTCTAAGTTTCTGTTGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 411
Db 37470 AAGTGAATCTGGCTTCTTATTTATGCTGCTTATTAAGAT---CAGCCTTACTCCGT 37414
QY 412 GTGTGTCATAGGACATAGATGATTTATTTCAATTTCTTTTATCTGTCGAGACTTGTT 471
Db 37413 GGTGATCTGATAGGTGATGATGATTTTAAATTTTGTATCTGTGAGCCCTGTTT 37354
QY 472 TGTTTTGAATATGTAATCAATTTTGGAG--AGTTTCTATAGGTCCTGACAGAAGGTAC 529
Db 37353 TGTGATCAATTTATGTTCAATTTTGGAGAGGTACCATGAGGTGCTGAGAGAAGGTAT 37294
QY 530 AGTC-TTTCTGTTTGTGTAATAGTCTGTAATAATCT-CTAGGTCCACTCTGTTTATGA 587
Db 37293 ATCTTTTGTTTTGTAGATTAATGTTTGTAGATATCTGTAGATCCATTTGTTTCATAA 37234
QY 588 CATCAGTTAGCTCCAGCAATTTCTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCT 647
Db 37233 CTCTCTGTAGTTTCACTGTGCTCTGTTTGTAG-TTTCTGTTTCCAGGATCTGTCCATGTT 37175
QY 648 GAGAAATGGGGTATTGAGTAGTACCCCACTATCTGTGTGTGAGGT-CAATATGTGATTTTAC 706

```


Db 37174 GAAAGTGGGCTGCTGAAAGTCTCCCACTATTATGTTGGAAGTGAATGTTGCTTTGAGC 37115
Qy 707 TGTAGCTGTCTGTTTATGAACCTGGGTGACATTTGTTTGGTGCATAGACATTAAGA 766
Db 37114 TTTACTAAAGTTTCTTTAATGAATGCTGCTGCAITTTGCAITTTGGAGCATATATCAGA 37055
Qy 767 ATTGCAATGCTCTTGTGTGAATTTT---CCTTGAAGCTCATGATGATTTCTCCCAATC 824
Db 37054 ATTGAGAGTTTCATCTGTGTAGATTTTACCTTTGATGATTAATGAAGTGTCTCTCTCTTT 36995
Qy 825 TCATCTGCTTAGTTTGGGTTTAAAGTCTATTAAGTCAAGATATTAATAAGTCAATGATCGGCT 884
Db 36994 TTGATAACTTTGGGTTGGAAGTCAAT---TTTATTTGATTAATGAAGTCTTCCAGCT 36938
Qy 895 TGTCTTTAGGGCCATTTGCTTAGAATA-TCTTTCCATCTCTTTTACTCTTAAGTGAATG 943
Db 36937 TGTCTTTCTATACCACTTTGCTTTGGAATAATGTTTCCAGCCCTTTTACTCTGAGTAGTGT 36878
Qy 944 CTAATCCAT---GGTAGGTTGTCTTTTGGATGAGAGTAGGATGATCTGTTGTTTCAT 1000
Db 36877 CTGCTTTATCCCTGAGTGGGTTTCTGTAAAGCAGCAAAATGTTGGTCTCTGTTGTGT 36818
Qy 1001 ATCCATCTGTATCCCAAGTATCTTTTCTAGAGAAATTAAGATCAITTAGTCAATGATGT 1060
Db 36817 ACCCAGTTGTTAGTCTATGCTTTTCTTGGGAATTTGAG-----TCCATTTGATGT 36766
Qy 1061 TGAGAAATTAATCAATGAGCAGTGTGTTGGATCTTGTATCTTGCATCTGGAAGTGTGT 1120
Db 36765 TAAGAGAAATCAAGAAAAGTAATGTTGTTCTCTGTCAITTTTGTGTTAAAGTTGGGA 36706
Qy 1121 GT 1180
Db 36705 ATCTGTTCTTATGCTGTCTTCTTTAGTTTG----- 36673
Qy 1181 GT 1240
Db 36672 -----TTAAAGGAT 36664
Qy 1241 TTATTAATCATATTTCTTGAATGGGTACATCTTTAGATGAAGTCTTTTCCCTAGC 1300
Db 36663 AACCTTCTGCTTTTCTAGGGTGAATTTCCATCTTGTGTGTGTGTGTGTGTGTGTGTGT 36604
Qy 1301 CTTCTTT---AGGCTGCAATTTGAAGATAGATATTTTACATCTGATTTTATCTTAGAAT 1358
Db 36603 ATCCTTTGAAGGCTGGATTCGTGGAAGATATTCGTGAATTTGGTTTGTCTCATGAAAT 36544
Qy 1359 GTCTTTCTTTCCCACTATTGTGACAGAAAGTTTCTTAAGTGCAGTAGTCTGGCTGTA 1418
Db 36543 ACTTTGATTTCTCCATCTATGCTAATGAGAGTTTGGCTGGGTATAGTCTGGCTGG 36484
Qy 1419 CATCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGAGGCTCTTACATTTTGTAGTT 1478
Db 36483 AATTTGTGTTCTCTTAGTGTCTGTATAACATCTGTCCAGGATCTTCTGGCTTTCTAGTC 36424
Qy 1479 TCTATTTGAAAGTCAAGTGTAAATCTTAATACATCTGCTTTTATATGTTAATTTGCTTTT 1538
Db 36423 TCT---GGAGAGAAGTCTGGAGTAATTTCTAATAGGCTGCTTTATATGTTACTT---GACCTT 36366
Qy 1539 TTTCCCTTGCATCTTTTAATATCTTTCTTTTGTCTATCTTTTGTAGTGAATTTGATTA 1598
Db 36365 TTTCCCTTACTGCTTATAATATCTATCTTTTATTTAGTGAATTTATTTGTTCTGTAATTA 36306
Qy 1599 TGCACTGTGGGAGTTTCTTTTCCGGTCCAACTATTTTGTGTGTTTGTATGCTTCTTTGTA 1658
Db 36305 TGTGTGGGAGGAATTTCTTTTCTGGTCCAGTCTATTGAGAGTACTGTAGGCTTCTTATA 36246
Qy 1659 CTTTGAAGGATCTCTTTCTCAAGGTAGAAATTTTCTTTTGTGTTTCTTGAATA 1718
Db 36245 TGTTCATGGGCACTCTCTTTCTTTTAGGTATGGAAATTTTCTTCTATAATTTTGTGAAGA 36186
Qy 1719 TATTTTCCCTGCTTTTGACCT-----GCCCTCTCCCTTCCCTTATCTCTTT 1766

Db 36185 TATTTGCTGGCCCTTTTAAAGTTAAATCTTCATCTCACTCTCTATTAATCCGTAGTTT 36126
Qy 1767 GGTTTTTCATAGTGTCTCTGGCTTCTCTGGATCTTTTATGCTCTGGATTTTATAGACTTA 1826
Db 36125 GGTCTTCTCATTTGTCCTGATTTCTCTGGATCTTTTGTAGTAGGATCTTTTGTCAATTC 36066
Qy 1827 ACATTTTCTTTGACCAAGGATATCCATTTCTTCTATCTGTCTTCACTGCTGTGATCTC 1886
Db 36065 ACATTTTCTATGATATATATCCATCTCCCTATGGAATCTTCTGCACCTGAGATCTC 36006
Qy 1887 TCTTCTATCTCTTGTATTTCTGTCAGTGAGCTTGTCTCTGAGGTTCTCTG-TTGGGTTCTT 1945
Db 36005 TCTTCCATCTCTTGTATTTCTGTGCTGATGCTCACATCTATGTTTCTGATTTCTTCTCT 35946
Qy 1946 AATTTTTCATTTTCCAGATTTCTCTCAGTTTGGGTTTCTTTTATTAATTTCTATTTCCACT 2005
Db 35945 AGGTTTCTATCTCCAGAGTTGCTTCTTGGGTTTCTTCTATTTCTTCTTCTTCTTCT 35886
Qy 2006 TCCAGTCTCGAATGTTTCTTACTCATTTTCTCT---CCCAGTATTTACATTTTCATAGTT 2062
Db 35885 TTTAGATCTTGGATGTTTGTTCATTCATTCGCTGCTGTTGTTGTTCTGTTTCTCTTAAT 35826
Qy 2063 TCTTTAATGATTTTATTCATTTCTCTCTTC 2091
Db 35825 TCTTTAAGGATTTTGTGTTTCTTCTTCTAC 35797

RESULT 26

US-10-087-192-1033
; Sequence 1033, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1033
; LENGTH: 200400
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)---(200400)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1033

Query Match 25.2%; Score 573.4; DB 1; Length 200400;
Best Local Similarity 63.4%; Pred. No. 1.5;
Matches 1164; Conservative 0; Mismatches 546; Indels 126; Gaps 20;
Qy 292 TCTATTTCTTGAATTTCTATCTTTGGCTCATTTTAACTCAGTAGTGTGTTGGTTTCC 351
Db 157315 TTTCTTTCTTAATTTCTTACCTGACTAATTT-----AGAGTAGAGAGTTGTTCAAGTTCC 157369

Qy 352 ATAAAGTTTCTGAATTTCTGTTGTTTCTGTTGTTGTTGTTTATCTAGATTTAAAGCTGT 411
Db 157370 ACGTGATCTGGGCTTTTGTGCTGTTTGTCTGTTATTAATGA---CCAGCCTTAAGCCAT 157426
Qy 412 GGTGTGCATAGGACATAGATATTATTTCAATGTCTTTTATCTGTGCGAGCTTGTCT 471
Db 157427 GGTGATCTCATAGGATGATGGGATTTATTTCTGTTCTTGTATCTGTTTAGGTTGTTT 157486
Qy 472 TGTTTGAATATGATTTCAATTTTGGAG--AGTTTTCATAGGTTGCTGACAGAGGTAC 529

RESULT 28
US-10-087-192-2029
; Sequence 2029, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01

Db 1052 AGAGTGGGCTGCTGAAATCTCCCACTATTATTGTATCGGGTGTGATGTGCTTTGAGAT 9993
Qy 708 GTAGCTGTCTGTTTATGAAGCTGGGTGACATGTGTGTGGTGCATAGACATTAAGAA 767
Db 9992 TTAGTAAAGTTTCTTTTATGAATGTGGTGCCTTGCATTTGTAGCATATGTTTCAGAA 9933
Qy 768 TTGCAATGCTCTCTGTGGGAT-TTCTTTGATGCTCTAGTATCTTCCCAATCTC 826
Db 9932 TTGAGAGTTCATCTTGGCAGATGTTTCTTTGACAGTATGAAGTCTCTTCTTATCTT 9873
Qy 827 ATCTGCTTAGTTTGGGTTT-----AAGTCTATTAGTTCAGATTAATAAATGACTGTATC 880
Db 9872 TTTTTTTGTGATAAATCTGTTGAGAGTGAATTTATTCATATTAGAATGGCTACTCC 9813
Qy 891 GCGTGTCTCTTAGS-----GCCATTTGCTTAGAATA-TCTTTTCCATCTCTTTACTCTA 934
Db 9812 AGCTTGTCTTGGGAAACAAACCAATTTGCTTGGAAATTTGTTTCCAACTTGAACCTG 9753
Qy 935 AGGTGATCTCTATCCATG---GTAGGTGTCTTTTGGATGACAGCAGTAGGATGATCT 991
Db 9752 AGGTAGTCTGCTTGTCTGCTGCTGAGTGCATTTCTGTATGACAGCAAATGCTGGTCC 9693
Qy 992 TGTTCATATCCATCTGTTTACCCAGTATCTTTTCTAGAGAAATTAAGATCAATGAGT 1051
Db 9692 TGCCTACACATCCAGTCTGTAGTCTATGTCTCTTTTGGAGAAATGA-----GTC 9642
Qy 1052 CATTCATGCTGAGATTAATCAATGACGAGTCTTTGTGATCTTGTGA-TCTTGCACCTG 1110
Db 9641 CATTCATGCTTGAAGATATTGAAGAAAGTGAATTTACTTCTCTGTTATTTTGTGTTG 9582
Qy 1111 TGAAGT 1170
Db 9581 TTAGAGGTGGAATATGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9537
Qy 1171 GT 1230
Db 9536 ----- 9537
Qy 1231 CTGGAATATTATTATTATTTTCTTGAATGTGGTAAACATCTTTAGATTGAAGTTT 1290
Db 9536 -TTGAAAGATTACTTCTTCTGCTTCTTCTAGGGTGTAGTTTCCCTCTTGTGTGTGT 9478
Qy 1291 TTTCTCCTA--GCCTCTTTAGTCTGCAATTTGAAGATAGATATTTTACATCTGATTT 1348
Db 9477 TCCATCTATTATCTTTTGTAGAGTGGATTTGTGGAATGATATTGTGTAATTTGGTTT 9418
Qy 1349 ATCTTAGAATGCTTTCTTCTTCTCAACTATTGTGACAGAAAGTTTCTTAAGTGCAGTAG 1408
Db 9417 GTCATGGAATATCTTGGTTTCTCCATCTATGATTAATTCAGAGTTTGTCTGGGTATAGTAG 9358
Qy 1409 TCTGGCCTGACATCTGATCTCTTGTAGTCTGTAGACATCTGTGAGGSCCTTCTTAC 1468
Db 9357 CTTGGGCTGGCATTTGTGTTCTTCTTGGGTCTGTATGACATCTGCTAGGATCTTCTAGC 9298
Qy 1468 ATTTGAGTTCTATTGGAAGTCAAGTGAATTAATCTAATACATCTGCTTTATATCTTA 1528
Db 9297 TTTCAATGCTCTGGTGAAGTCTAGTCTTATCTGATAGGCTGCTCTTTATATGTTA 9238
Qy 1529 ATTTGCTTTTTTCCCTTGTGATCTTTTAAATATCTTTCTTTGTCTCTATCTTTTGTATGAT 1588
Db 9237 CTT-GACTTCTTCCCTTACTGCTTTTCAATTTATTTCTTTTGTGTTAATGCAATTTGATTT 9179
Qy 1589 TGTATATTATGCACTGTGGGAGTTCTTTTCCGTCCTAATCTATTGTTGGTCTTTGTAT 1648
Db 9178 TCAATTATTATGTGACAGAGGA-ATTCTTTTCTGTCTTCTTGTAGTATGGAATTTCTTCTAAT 9120
Qy 1649 GCTTCTGTGATCTGATAGGATCTTTCTTCTCAAGTTTAGGAAATTTTCTTTTGTGTT 1708
Db 9119 GTTTCTGTATGTTCAATGACATCGCTTTCTTTAGTATGGAATTTCTTCTTCTAAT 9060
Qy 1709 TTTCTGAAAAATTTTCCCTGCTTTTGTACCT-GCCTTCTTCCCTCTCTCTCTCTCTCTCT 1763
Db 9059 TTGTTGAAGATATTACTTGGGCTTTGTAGTTGGGAATCTTCACTTTCTTCTATACCTATT 9000

Qy 1764 -----TTTGGTTTTTGCATAGTGTCTCTGGCTTCTCTGATCTTTTATGCTCTGGATT 1814
Db 8999 ATCCTTAGGTTTCAATTTCTCTCAITGTGCTCTATATTTCTCTGATTTTGGTTAGGAGC 8940
Qy 1815 AATTTAGACTTAACATTTTCTTTGACCAAGGTATCCATTTCTTCTATCTTCTCTTCACTG 1874
Db 8939 TTTTTCATTTTGTATTTCTTTGACTGTGTGTCAATGTTTCTATGATATCTTCTGCA 8880
Qy 1875 COTGAGATCTCTCTCTATCTCTTGTATCTGTGAGTGTGTGTGTGTGTGTGTGTGTGT 1934
Db 8879 COTGAGATCTCTCTCTATCTCTTGTATCTGTGTGTGTGTGTGTGTGTGTGTGTGT 8820
Qy 1935 G-TTGGTTTCTTAAATTTTTCATTTCCAGATTTCTTCAAGTTTGGTTTGTATTAAAT 1993
Db 8819 GATCTTTTCTTAGGTTTCTGTCTCCAGCGTGTCTCCCTTTGTGATTTCTTATTGTT 8760
Qy 1994 TCTATTCCACTTTTCAAGTCTCTGAAATGTTTCTTCTCTCTCTCTCTCTCTCTCTCT 2050
Db 8759 TCTAGTTCCATTTTGTAGATCTCTGATGTTTGTCTCAITTTCTCTTCCCTGTTTGT 8700
Qy 2051 TTTTCTAGGTTTCTTAAATGATTTTCAATTTCTCTCTCTCTCTCTCTCTCTCTCT 2109
Db 8699 TTTCTCTGTAGTTCTTTAAAGATTTTGTGTCTCTCTTTAAGGCTTCTAGTTGTTT 8641

RESULT 34

US-10-087-192-1861/c
; Sequence 1861, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Engelhard, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1861
; LENGTH: 175338
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(175338)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1861

Query Match 24.8%; Score 564.9; DB 1; Length 175338;

Best Local Similarity 63.1%; Pred. No. 1.7;
Matches 1159; Conservative 0; Mismatches 551; Indels 127; Gaps 20;

Qy 292 TCTATTTCTTGTATTTCTATCTTGGCTCAATTTTAACTAGTAGTGTGTGTGTGTGT 351
Db 53933 TTTCTTTCTTTATTTTCAATTTTGACCAATATATCAATTAAGTAGAGCAITGTCAGCTTCC 53874
Qy 352 ATAAGTTCTTGAAGTTTCTGTTTCTGTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTT 411
Db 53873 ATGTGATGT 53817
Qy 412 GTGTGTGATAGGACATAGAGTATTTTCAATTTGTCTTTTATCTGTGAGACTTGTCT 471
Db 53816 GGTGATCTCATAGGGTGCATGCGATTAATTTCAATTTCTTTATATCTGTGTAGGCTGATT 53757
Qy 472 TGTTTTGAATATGATTTCAATTTTGGAGA--GTTTCTATAGGTTCTGTGACAGAGGTAC 529
Db 53756 TGTGACCAATTAATGTTGTTCAATTTTGGAGAAGGTACCATGAGTGTCTGAAAGAGGTAT 53697

QY	530	AGTCCTTTGGTFTTTG-GTGAAATAGTCTGTAAATATCT-CTAGGTCCACTTGGTTTATGA	587
Db	53696	ATTCTTTTGGCTTTAGAAATGTTCTATAAATATCTGTTAGATCCAAATTTGGTTTCATAG	53637
QY	588	CATCAGTTAGCTCCAGCATTTCTCTGTTTCGTTTTTTTGTGTAGATGACCTAACTCTTTGGA	647
Db	53636	CATCTGTTACTTTCACCTGCTCTCTGTTTAG--TTTGTGTTTCCATGATCTGTCCATGCT	53578
QY	648	GAGAAATGGGGTATTGAAGTAGCCCACTATCTGTGTGTAGAGT-CAATAATGTGATTTTATG	706
Db	53577	GAGAGTAGGGTGTGTGAAGTCTCCCACTATTAATTGTGTGTGTGTCAAGGTGTGCTTTTGAGC	53518
QY	707	TGTAGCTGTCTTGTGTTTTTAAGACTTGGGTGCATTTGCTTTTGTGTGCATAGACATTAAGA	766
Db	53517	TTTAGTAAAGTTTCTTTTATGAATGTGGGTGCCCTTGCATTTTGGAGTATAGATGTTCA	53458
QY	767	ATTGCAATGTCCTCTTGGTGGGA-TTTTCTTTTGATGCCCTATGTAGTATCTTCTCCAAATCT	825
Db	53457	ATTTAGAATGCATCTTTGGTAGATTTTCTCTTTGATGAGAAATCAAGATATCTCTTCTATTTC	53398
QY	826	CA---TCTGCTTAGTTTGGGTTTAAGTCTATTAGTCAGATATTAATAATGACTGTAT---	879
Db	53397	TATTTTGTGAYAACCTTTGGTGGAAAGTTGATTTTATGTACTAGAAATGAATGGATACT	53338
QY	880	-CGCGTGTCTCTTAGGGCCATTTGCTTAGAATA-TCTTTTCCATCTCTTTACTCTAAGG	937
Db	53337	CCAGCTTGTCTCTTGGGACATTTGCTTGGAAAAATGTTTTCAAGCCCTTACTCTGAGG	53278
QY	938	-----TGATGCTATCCATGTGTAGTGTCTTTTTCGATGCACGTAGATGGATCT	991
Db	53277	TAGTACTGTCTGCTTTTGATCTGAGGTGCATTTCTTGATGCACAAAAATGCTGGGTCC	53218
QY	992	TGTTTTCAATCCATCTCTTACCAGTATCTTTTTCTAGAGAAATTAAGATCATTTGAGT	1051
Db	53217	TACTTACATATCAAGTCTGTATGTCATGCTCTTTTATTAGAGAATTCAG-----TC	53166
QY	1052	CATTGATGTGAAGATTATCAATGACGAGTGTCTGGATCTTGTTATCTGCACTGT	1111
Db	53165	CTTTGAATTTAGAGATATTAAGGAAAAGTGATGTACTTCCATATTATTTGTGTGTTA	53106
QY	1112	GAAGTG	1171
Db	53105	GAGGTGGAATTAATGTTGTGTGGCTACCTTCTTTTGGGTTGT-----	53063
QY	1172	TGTGTGCTGTCTGTGTCTGTGTGTGTGTGTCTCCCCCTTTTGTATTTTGGCC	1231
Db	53062	-----	53062
QY	1232	TGGAATTAATTTATTTATTCATATTTTCTTGAATGGGTAAACATCTTTAGATTGAATTTT	1291
Db	53061	GAAGAAGATACCTTCTTGCTTTTCTAGGGTGTAGTTCCCTCCTGTGTGGATTTT	53002
QY	1292	TCTCCTAGCCTCTT--TAGGTCTGATTTGAAGATAGATATTCTTTACATCTCAATTTTA	1349
Db	53001	TTGCTGTATCTTTTGTAGGGCTGGAATTTGTGAAAGATATGTGTAAATTTGGTTTTG	52942
QY	1350	TCTTAGAATGTCTTTCTTCTCCAACTATTGTGACAGAAAGTTTTTCTTCAAGTCAGTAGT	1409
Db	52941	TCAATGAAATATAATGTTTTCACTGCTATGGTAAATTAGAGTTTGTCTGGTATAGTAGC	52882
QY	1410	CTGGCCTGACATCTGTAGTCTCTTGGAGTCTGTAGCAATCTGTGCAAGGCCCTTCTTACA	1469
Db	52881	CTGGACTGATATAATGTGTTTCTCTCAGGGTCTGTATGAGATCTGCTTAAGATCTTCTAGCT	52822
QY	1470	TTTTAGATTCTTATTGGAAAAAGTCAGGTGTAAATCTTAATACATCTGCCCTTTATATGTTAA	1529
Db	52821	TTCAACAGTCTCT-TGTGAGAAGTCTGGTTTAATCTCTGAGAGGTCTGCCCTTTATATGTTAC	52763
QY	1530	TTGGTCTTTTTCCTTGGCACTTTTAAATATCTTTCTTTTGTCTATACTTTTAGTGATT	1589
Db	52762	TT-GACCTTTTCCCTTACTGCTTTTAAATATCCTTCTTTGTTTTGTGCATTTGGTGTTT	52704

Qy	1590	TGATTATATGCACTGTGGGAGTTCCTTTCCGGTCCAACTATTTGGTGTGTTGTATG	1649
Db	52703	TGATTATATGTGATGGGAGAAATTTCTTTCTGGTCCAGTCTATTTGGAGTTCGTAGG	52644
Qy	1650	CTTCTTGCTACCTTGATAGGACATCTCTTCTCAAGGTTAGGAATTTTCTCTTT	1705
Db	52643	CTTCTTGATGTTTCATGGGCAICTCTTCTTAGTTAGGTAGGGAAGTTTCTCTACACTT	52584
Qy	1706	GTTTCTTGAAAAATATTTCCCTGCTTTTGACCTGCTCTTCTCCCTTCC-----TCT	1758
Db	52583	TTTTTTGTTGAAGATAATTTACTGACCCCTTTAAGTTGGTAAATTTCTATACCCTAATTATCC	52524
Qy	1759	ATTCCCTTGGTTTTTCATAGTGTCTCTGGCTCTCTGGATCTCTGATGTTTATGCCGTGATTATTT	1818
Db	52523	TTAGATTTGGTCTTCCCATGTGTCTCGAATTTCCCTGGATGTTTGGAGTTAGGAGCTTTT	52464
Qy	1819	TAGACITTAACATTTTCTTTGACCAAGATATCCATTTCTCTATCTTGCTTTCACCTGCCCTG	1878
Db	52463	TGCATGTTGTGTTTTCTTTGACCTGTCTGTGATGTTTTTATGCTATTTTCTGCACCTG	52404
Qy	1879	AGATTCCTCTTCTATCTCTGTATCTGTCACTGAGGCTGTCTCTGAGGTTTCTCTG	1937
Db	52403	AGATTCCTCTCTCTATCTCTGTATCTGTGCGCATGCTTGCACTATAAAGTCTCTGATC	52344
Qy	1938	GGGTTCCTTAAATTTTTTCATTTCCAGATTTCCCTTCAGTTTGGGTGTTTGTATTAATTCCTA	1997
Db	52343	TCCTTCTAGGTTTTCTAACTCCAGGCTGTCTCTCTTTTGTATCTCTTTATTTGTTTCTA	52284
Qy	1998	TTTCCACTTTCAGGTCCTGAAATGTTTACTCATTTTCT---CCAGATTTTACATTTT	2054
Db	52283	TTTCTCTTTTAGATCTTGATGGTGTGTTGTTTCATTTCTTTTGCCCTATTGATTTGTGTTT	52224
Qy	2055	CATAGGTTTCTTTAATGGAATTTATTCATTTCTCTCTC	2091
Db	52223	CCGTGCATCTTTTAGGGAATTTGTGTGTGTGTTTC	52187
RESULT 35			
US-10-322-281-495			
; Sequence 495, Application US/10322281			
; Publication No. US20040126762A1			
; GENERAL INFORMATION:			
; APPLICANT: David W. Morris			
; APPLICANT: Marc S. Malandro			
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer			
; FILE REFERENCES: 529452001000			
; CURRENT APPLICATION NUMBER: US/10/322,281			
; CURRENT FILING DATE: 2002-12-17			
; NUMBER OF SEQ ID NOS: 866			
; SOFTWARE: FastSeq for Windows Version 4.0.			
; SEQ ID NO 495			
; LENGTH: 176594			
; TYPE: DNA			
; ORGANISM: Mus musculus			
; FEATURE:			
; NAME/KEY: misc_feature			
; LOCATION: (1)...(176594)			
; OTHER INFORMATION: n = A,T,C or G			
US-10-322-281-495			

RESULT 35

```

US-10-322-281-495
; Sequence 495, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0.
; SEQ ID NO 495
; LENGTH: 176594
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(176594)
; OTHER INFORMATION: n = A,T,C or G
US-10-322-281-495

```

	Query Match	24.5%;	Score 558.6;	DB 1;	Length 176594;
	Best Local Similarity	62.5%;	Pred. No. 1.8;		
	Matches 1149;	Conservative	0;	Mismatches 559;	Indels 130; Gaps 19;
QY	292	TCTATTTCTTGATTTCTATCTTTGGCTCATTTTAACTCAGTAGTAGTGTGTTGGTTTCC	351		
Db	129540	TTCTTTCTTTATTCCTTCCTTGACCAAGGTATCATTGAAAGAGTGTGTTCCCGTTTCC	129599		
QY	352	ATAAGTTGTGAAGTTTCTGTTTCTCGTTGTTGTTGTTGTTATCTAGATTTTAACTGT	411		
Db	129600	AGCTGAATGTGGCTTCCATATCTATGTTGTTATTGAAGATC---AGCCTTAGTGCAT	129656		
QY	412	GGTGTCAGATAGGACATAGAGTATATTTCATTTGCTTTTATCTGTCGAGACTTGCTT	471		

Db 129657 GGTGATCTGATAGGATCGATGGACAAATTCCTGTTGATCTGTTGAGGCGCTGTTT 129716
Qy 472 TGTGTTGAAATATGATTCATATTTGGAGA--GTTTCATAGGTCCTGACAGAGGTAC 529
Db 129717 TGTGACCAATATATGTTGCAATTTGGAGAGGTACCATGAGGTGCTGAGAGAAATGAT 129776
Qy 530 AGTCTTTGTTGG--TGAATAGTCGTGAAATATCT-CTAGGTCACATTTGTTTATGA 587
Db 129777 ATCCCTTTGGTTAGGATAAAGTTCGTGATATCTGTTCAATCTATTTGTTTATAA 129836
Qy 588 CATCAGTTAGTCCAGCAATTTCTCTGTTTCGTTTGTGAGATGACCTAACTGTTGGA 647
Db 129837 CTTCTGTTAGTTTCACTGTGCTGCTGTTAG--TTTCTGTTTCCATGATCTGTTCAATGGT 129895
Qy 648 GAGATGGGTATTGAGTACCCACATCTGCTGTGAGGT-CAATATGTAATTTTACG 706
Db 129896 GAGATGGGTGTTGAAATCTCCACATTAATTTGTTGAGATGCAATGTTGCTTTGAAC 129955
Qy 707 TGTAGCTGTCTGTTTATGAATCTGGGTGACATTTGTTGTTGCTGATAGACATTAAGA 766
Db 129956 TTTACTAAAGTTTCTTTAAATAATGTTGCTGCCCTTGCAATTTGAGCATAGATATTGCA 130015
Qy 767 ATTGCAATGCTCTGTTGAGATTTT-CCTTTGATGCTATGATGATTTCTTCCCAATCT 825
Db 130016 ATTGAGATTTCTCTTGAAGATTTTACCTTTGATGATGAAAGTGTCCCTCTCTGCT 130075
Qy 826 CATCTGCTTAGTTTGGGTTTAAAGTCTA-TTATGTCAGATATTAATAATGACTGTATCGGCT 884
Db 130076 TTTTGTATAACTTTTGGTTGAGTCTGATTTTATCAATATTAGATGGCTACTCCACT 130135
Qy 885 TCTTCTTAGGCGCAATTTGCTTGAATA-TCTTTTCAATCTTTTACTCTTAAGTGATGT 943
Db 130136 TGTCTTCTCAGGCCAATTTGCTTGAATAATGTTTCCAGCCTTTCACTTCTGAGGTAGTGT 130195
Qy 944 CTATCCAT--GGTAGGTGCTTTTGGATGAGAGTAGGATGGATCTGTTTTCAT 1000
Db 130196 CTGCTTTTCCCTGGGATGGGTTTCTGTAAGCAGCAAAATGTTGGTCTGTTGTTGT 130255
Qy 1001 ATCCATCTGTATCCAGATATTTTCTAGAGAAATTAAGATCATTTAGATGATGATGT 1060
Db 130256 AGCCAGTCTGTAGTCTATGCTTTTATGCGGAATTGAG-----TCCATTTGATAT 130307
Qy 1061 TGAGATTAATCAATGAGCAGTGTGTTGATCTGTTATCTTGCACCTTGTGAGTGTGT 1120
Db 130308 TAAGAGATTAATAGGAAAGTAATGTTGCTTCTTATTTTGTGTTTAAAGTTGGCA 130367
Qy 1121 GTGCTGT 1180
Db 130368 TTCCTGTTGTGGCTGCTCTTTTAGTGTG----- 130400
Qy 1181 GTGCTGT 1240
Db 130401 -----TTGAAGAT 130409
Qy 1241 TTATTATTCAATTTTCTGAAATGGGTAAATCTTTAGATTGAAATTTTCTCTTACG 1300
Db 130410 TACTTCTGCTTTTCTAGGTCGTGTTTCACTTTTATGTTTATTTTCTTCTTCT 130469
Qy 1301 CTCTTT-----TAGGTCTGATTTGAAGATAGATTTCTTTACATCTGATTTTATCT 1352
Db 130470 GTTGTATCTCTTGAAGGCTGATTCATCGAAGATAATGTGCAATTTGCTTTGTCG 130529
Qy 1353 TGAATGCTTCTTCTTCCAACTATTGTGACAGAAAGTTTCTTAAGTCAGTGTGCTG 1412
Db 130530 TGAATACCTTTGGTTCTCCATCTAATGTAATGAAAGTTTGGCTGGGTATAGTACCTG 130589
Qy 1413 GCCTGACATCTGATCTCTTGGAGTCTGTAGACATCTGTCCAGGCGCTTCTTACATTT 1472
Db 130590 GCCTGGCAATTTGTTCTTCTAGTCTGTATTAACATCTGTCCAGGCTCTTCTGGCTTC 130649
Qy 1473 TGAGTTTCTATTGGAAAGTCAGGTGTAATCTATACATCTGCTTTATATGTTAATG 1532

Db 130650 ATAGTCTCTGGT-GAAAACTGTGGTGTGATTCATAAGCCCTGCCCTTATATGTTACTTG 130708
Qy 1533 GTCCTTTTCCCTTGCATCTTTTAAATATTTCTTTTCTTATATCTTTTCTTTTCTTTTGA 1592
Db 130709 ACCCTTTTCTC--TTATGCTTTTAAATATTTATCTTTATT-TGGTGTGTTTTTGTCTGA 130766
Qy 1593 TTATTATGACACTGTGGGAGTTCTTTTCCGGTCCAACTATTGCTGTTTGGTGTGTTGATGCTT 1652
Db 130767 TTATTATGCTGTGAGGAGAAATTTCTTTCTAGTCCAGTCTATTGGAGTTCTGTAGGCTT 130826
Qy 1653 CTTGTACCTTTGATAGGATCTCTTTCTCAAGGTAGGAAATTTTCTTTTCTTTTGGTTTCT 1712
Db 130827 CTTGTATGTTCAAGGATCTCTTTCTTTAGGATGGGAAGTTTCTTCTATATTTTGT 130886
Qy 1713 TGAATAATTTTCCCTGCTTTTGACCTG-----CCTTCTTCCCTTCTCTCT 1758
Db 130887 TGAAGATATTGCTGGCCCTTTTAAAGTTGAAATCTTTCTTCTGCTACTCTATATAT 130946
Qy 1759 ATCTCTTGGTTTTCATAGTGTCTCTGGCTTCTGGATGTTTATGCTGGATTTAT 1818
Db 130947 GTAGGTTTGTCTCTCATTTGCTCTAAATTTCTTAGAAGTTTGTACTCAGATCTTTT 131006
Qy 1819 TAGACTTAACTTTCTTTTGACCAAGGTATCCATTTCTTATCTTGTCTGCTCACTGCTG 1876
Db 131007 TGCAITTTGCAITTTTCTTTGATTGTTGCGGATGTTCTATGTAATCTTCTGCACCTG 131066
Qy 1879 AGATCTCTCTTCTATCTTGTATTTCTGTAGTCAAGGTTGTCTCTGAGGTTCCCTG--T 1936
Db 131067 AGATTTCTCTCTTCCATCTCTTGTATTTGTTGCTAATGCTGCATCTATGTTCTAGATT 131126
Qy 1937 TGGGTTCTTAATTTTCTCAATTTCCAGATTTCTTCAAGTTGGGTTTGTATTAAATCT 1996
Db 131127 TTTTCCCTAGGTTTCTATCTCAGGTTGCTCATTTGGGTTTCTTATTTGTGCT 131186
Qy 1997 ATTTCCACTTTCAGGTCCTGAAATGTTTATCTCATTT--TCTCCAGTATTTTACATTT 2054
Db 131187 ACTTCACTTTTATGTTCTGTTATGATTTTGTTCATTTCCACCCTGTTTGGTTGTGTTT 131246
Qy 2055 CATAGTTTCTTAAATGGAATTTTCAATTTCTCTCTTCA 2092
Db 131247 TCTGTTTTTCTTAAGATTGTTACTGTACTCTTTTA 131284

RESULT 36

US-10-175-523-75
; Sequence 75, Application US/10175523
; Publication No. US20030396264A1
; GENERAL INFORMATION:
; APPLICANT: Brockman, Jeffrey
; APPLICANT: Evans, David
; APPLICANT: Hook, Derek
; APPLICANT: Klimczak, Leszek
; APPLICANT: Laeng, Pascal
; APPLICANT: Palfreyman, Michael
; APPLICANT: Rajan, Prihi
; TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)
; FILE REFERENCE: 3235/14795-US3
; CURRENT APPLICATION NUMBER: US/10/175,523
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 60/299,151
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/317,828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/325,150
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/333,047
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 60/349,936
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/361,834
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1

31891	Db	-----TCTCTGTGATCTCTGGTGAGCGCTTGCATCTGACCTCTGATCTCTGCC	31942
1945	QY	TAATTTTTTCATTTCCAGATTTCCCTCAGTTTGGGTTTTGTTTATTAATTTCTATTTCCAC	2004
31943	Db	 TAGGTTTCTATCTCCGGTTTGTCTCCCTTTGTGATTTCTTTATTTGTTTCTATTTCCA-	32001
2005	QY	TTTCAGGTCCTGAATGTTTTTACTCATTTTCTCTCCC--AGTATTATCATTTTTCATAGGT	2061
32002	Db	TTTTAGATCTTGATAGTTTTTGTTCATTTTCCCTTCGCCGTTTGTGTTGGTTTTCTCTGAA	32061
2062	QY	TTCTTTAATGGAGTTTATTCATTTTCTCTTTCAGGACCTTTTATGAAATTCATAA	2114
32062	Db	TTCTTTAAGGAGATTTATGTCCTCTCTTTAAAGTCCCTATCATCATCATCA	32114

```

RESULT 38
US-10-087-192-1765
; Sequence 1765, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1765
; LENGTH: 68370
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(68370)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1765

```

	Query Match	24.2%;	Score 530.599;	DB 1;	Length 68370;
	Best Local Similarity	62.0%;	Pred. No. 2.9;		
	Matches 1153;	Conservative 0;	Mismatches 584;	Indels 124;	Gaps 19;
Qy	292	TCTATTTCCTGATTCTCATCTCTGGCTCATTTTA	CTCAGTAGTGAGTTGTTGGTTTCC	351	
Db	45984	TTTCTTCTTTATTCATCCITTCGACCAAGGTATCA	TTGAGAAGAGTGTTATTCCTGTTTCC	46043	
Qy	352	ATAAGTTTGTAAAGTTTTCGTGTTTCGTGTTCTG	TGTTGTTGTTATCTAGATTTTAAGCTGT	411	
Db	46044	AGTTTAAAGTTGGCTTTCGAATATTTATGTTGTC	ATTCATTTGAAGATG---AGCCTTAGTCTAT	46100	
Qy	412	GGTGTCGATAGGACATAGAGTATTTATTCAAAT	TGCTTTTATCTGTCGAGACTTTGCTT	471	
Db	46101	GGTGGTCTCATAGGATGCATGGACAAATTTCAAT	ATTTTCTGTTAGGCGCTGTTT	46160	
Qy	472	TGTTTTGAAATATGATTTCAATTTTGGAGA--	GTTTTCATAGGGTCTGCACAAAGGTAC	529	
Db	46161	TATGACCAATATATGGTCAATTTTGGAGAAGGTC	CCCGTGAGGTCTCGAGAAGAGGTAT	46220	
Qy	530	AGTC-TTTCTGTTTTGGTGAANATGCTGCTAAATA	-TCTCTAGGTCCACTTGGTTTATGA	587	
Db	46221	ATCCTTTCTGTTTAGGATAAAATGTTCTGTAGAT	ACCCGTTAAGTCCATTTGTTTTCATAA	46280	
Qy	588	CATCAGTTAGCTCCAGCATTTCTCTGTTTCGTTT	TTTTTGTGAGATGACCTAACTGTTTGA	647	
Db	46281	CTTCGTTTAGTTTCACTGTGTCCTGTTTAG-TTT	CTGTTTTCCAGATCTCTCTATTGAT	46339	
Qy	648	GAGAATGGGGTATTGAAGT-AGCCCCATCATCTG	TGTGTGAGGT-CAATATATGTGATTTTAG	705	

QY 1764 TTGTTGTTTGCATAGTGTCTGGCTTCCTGGATGTTTTTANGCTGGATTTATTTAGAC 1823
DB 47392 TTGTTGTTTCTCATGTGTCTGGATTTCTGGATGTTTTGAGTTAGGATCCTTTTGCAT 47451
QY 1824 TTAAATTTTCTTGGACCAAGGTATCCATTTCTCTATCTTGTCTCTCACTGCTGAGATT 1883
DB 47452 TTGTTATTTTCTTGAATGTTGTCGGATGTTCTCTAGGAACTCTTGGACCTGAGATT 47511
QY 1884 CTCTCTCTATCTCTGTTATTTCTGTCAGTGAGGCTTGTCTGAGGTTCTTG-TTGGTT 1942
DB 47512 CTCTCTCTCAATTTCTGTTATTTGTTGTCGTGATGTCGAATCTATGTTCCAGATTCTCTT 47571
QY 1943 CTTAATTTTTCATTTCCAGATTCTCTGATTTGGGTTTGGTTTATTAATTTCTATTTCC 2002
DB 47572 CTTACGGTTCTATCTCCAGGCTGTTCTCGTTTGGGTTTCTTCTTATTTGTTCTACTTCC 47631
QY 2003 ACTTTCAGTCTGCTGAATGTTTACTCATTTTCTCTCCAGTATT---TACATTTTCTATAG 2059
DB 47632 CTTTCTAGTATGTTTGTTCATTTCCATCACCCTGTTGGATGTTTCTCTGT 47691
QY 2060 GTTCTTTAAAGATTTATTTCAATTTCTCTCAAGGACCTTTTATGAAATTCATAAATGT 2119
DB 47692 TTTTCTTTAATGACTTCTAACTGTTTGTGTTGTTTCTCTTCTTTTCTTTTAAAGACCTGT 47751
QY 2120 A 2120
DB 47752 A 47752

RESULT 39

US-10-175-523-97
; Sequence 97, Application US/10175523
; Publication No. US20030096264A1
; GENERAL INFORMATION:
; APPLICANT: Brockman, Jeffrey
; APPLICANT: Evans, David
; APPLICANT: Hook, Derek
; APPLICANT: Klimczak, Leszek
; APPLICANT: Laeng, Pascal
; APPLICANT: Palfreyman, Michael
; APPLICANT: Rajan, Prithi
; TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)
; FILE REFERENCE: 3235/IJ795-US3
; CURRENT APPLICATION NUMBER: US/10/175-523
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 60/299,151
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/317,828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/325,150
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/333,047
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 60/349,936
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/361,834
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 97
; LENGTH: 106664
; TYPE: DNA
; ORGANISM: Mus musculus domesticus
US-10-175-523-97

Query Match 24.0%; Score 547.299; DB 1; Length 106664;
Best Local Similarity 61.6%; Pred No. 2.5;
Matches 1139; Conservative 0; Mismatches 572; Indels 137; Gaps 18;

QY 292 TCTATTTCTGATTTCTATCTGGCTATCTTTTAACTCAGTAGTGTGTTGGTTCTCC 351
DB 93027 TTTCTTTCTTTATCTCTCTTGAACCAAGGTATCATTCAGAGAGAGTGTGTTCTATTTTC 93086

QY 352 ATAAGTTTGTAAAGTTTCTGTGTTCTGTTGTTGTTTATCTAGATTTTAAAGCTGT 411
DB 93087 ACATGAATCTGGCTTTCTTATTTATTTATTTGTTATTTGAAGAT---CAGCCTTAGTCCAT 93143
QY 412 GGTGGTCAGATAGGACATAGAGTATTTATTTCAATTTGTTCTTTATCTGTGAGACTTGTGT 471
DB 93144 GATGATCTTATAGGATACAGCGACAAATTTCAATATTTTGTATCTTTGAGGCTTGT 93203
QY 472 TGTGTTGAAATATGATTCAAATTTTGGAGA--GTTTCATAGGGTGTGACAAAGAGTAC 529
DB 93204 TGTGACCAATATATGTGTGTCAGTTTGGAGAAGGTACCATGAGTGTGTCGAGAAGAAAGTAT 93263
QY 530 AGTC-TTTCGTGTTTGGTGGAATAGTCTGTAATATCT-CTAGGTCCACTTGGTTTATGA 587
DB 93264 ATCCTTTGTGTTTGGATTAATAATGTTCTGTAGATATCTGTTAGATCCATTTGTTTCATAA 93323
QY 588 CATCAGTTAGTCCAGCATTTCTCTGTTTTCGTTTGTGTTGATGACCTAACTGTTGA 647
DB 93324 CTTCTGTTAATTTCACTGTGTCCCTGTTTGTAG-TTCTGTTTCCATGATCTGTCCATTTGT 93382
QY 648 GAGATGGGTTATTTGAAGTAGCCCACTATCTGTGTGAGGT-CAATATGATTTTATG 706
DB 93383 GAAAGTGTATGTGAAGTCTCCCACTATTTATTTGTGAGCTGCAATGTGCTTTGAGC 93442
QY 707 TGTAGCTGTGTTGTTTATGAACTTGGGTGACATTTGTTGGTGCATAGACATTAAGA 766
DB 93443 TTTTACTAAAGTGTCTTTAATGAATGTGCTGCCCTTTATTTGGAGCATAGATATTCAGA 93502
QY 767 ATGCAATGCTCTGTTGGGATTTTCC-TTGTAGTCCCTATGTAGTATTTCTTCCAAATCT 825
DB 93503 ATTGATAGTTCCTCTTGGAGGATTTTACTTTTGTAGTATGATGAAGTGGCCCTCTTGTCT 93562
QY 826 CATCTGCTTAGTTTGGGTTTAAAGTCTA-TTAGTCAGATATTTAAATGACTATCTCGCT 884
DB 93563 TTTTGTATGACTTTGGGTTTGAAGTGGATTTATTTAGACATTTAGAAATGGTACTCCAGCT 93622
QY 885 TGCTTCTTAGGGCCATTTGCTTTAGAAATA-TCTTTTCCATCTCTTTTACTCTAAGGTGATGT 943
DB 93623 TGTTTCTTCCAGACCAATTTGCTTTGGAAAATTTGTTTCCAGCCTTTTCAATTTGAGGTAGTGT 93682
QY 944 CTATCCAT---GGTAGTGTCTTTTGGATGCGACGATGAGTGGATCTTGTGTTTCTAT 1000
DB 93683 CTGTCTTTTCCCTGAGATGGGTTTCCGTAGACGACAAATTTGGGTCTCTGTTTCTGT 93742
QY 1001 ATCCATTTCTGTACCAGTATCTTTTCTAGAGAAATTAAGATCATTTGAGTCAATTTGATGT 1060
DB 93743 AGCCAGTCTTTTGTAGTCTATATTTTGGCAGAGTTGA-----GTCCATTTGATAT 93793
QY 1061 TGAGAAATATCAATGACGAGTGTGTTGGAATCTTGTATCTTGCACTGTGGAAGTGTGT 1120
DB 93794 TAAGAGATATTAAGGAAAAGTAAATTTGCTCTTCTTATTTTGTGTTAAAGTTGCA 93853
QY 1121 GT 1180
DB 93854 TTCTGTTCTGTGGCTGTCTCTTTTGTAGTTTG----- 93886
QY 1181 GT 1240
DB 93887 -----TTGAGGAT 93895
QY 1241 TTATTTATTCATATTTTCTTGAATGTGGTAAACATCTTTTAGATTGAAGTTTTTCTC---C 1296
DB 93896 TACCTTCTGCTTTTCTTAGGCGGTGTTTCCATCTTGTATTTGTTTTTCTGTTA 93955
QY 1297 TAGCCTTCTTTAGGTTCTGCATTTGAAGATAGATATTTCTTTACATCTGATTTTATCTTGA 1356
DB 93956 TTATCTTTTGAAGCGCGGATTTTGGAAAAGATATGTTAAATTTGGTCTTTTGTGTA 94015
QY 1357 ATGTCTTTCTTCTCAACTATTTGTGACAGAAAGTTTTTCTAAGTGCAGTACTCTGGCT 1416
DB 94016 ATACTTTTCTTCTCATCTAAGTAAATGAAAGTTGGCTCGGTACAGTACCTGGTT 94075

Db 43096 TCCTTTGAAGGGCTGGATTCTGTGGAAGATAATGTGTGAATTTGGTTTGTCTATGAATA 43155
QY 1360 TCCTTTCTTCTCCAACTATTGTGACAGAAAGTTTCTTAAGTGCAGTAGTCTGGCCGTGAC 1419
Db 43156 CTTTGGTTTCTCCATCTATGTAATGGAAG-TTTGCTGGGTATAGTACCTATGCTGGC 43214
QY 1420 ATCTGTAGTCTCTTGAGTCTGTAGACATCTGTGACGGGCTCTTCTACATTTTGTAGTTT 1479
Db 43215 ATTGTGTCTCTTAGTGTCTGTATATAACATCTGTCCAGGATCTTCCAGACTTTCATAGTCT 43274
QY 1480 CTATTGGAAGTCAAGTGTAACTTAATACATCTGCTCTTATATGTTAAATGTGCTTTT 1539
Db 43275 CTGGT-GAAAAGCTGGGTAAATCTGTATAGCTTGGCTTTTATTTTACTTGTACCTTTT 43333
QY 1540 TTCCCTTGATCTTTTAATATCTTCTTCTTCTTCTATATCTTTTGTAGTGAATTTGATTTAT 1599
Db 43334 TTCCCTTACTGCTTTTAATATCTATCTTTTATAGTGTACTTGTGTCTGATTTAT 43393
QY 1600 GCACGTGGGAGTTCTTTTCCGGTCCAACTATTGGTGTGTGTGTATGCTTCTGTAC 1659
Db 43394 GTGTCAAGGAGAAATCTTCTTCTGCTCCAGTCTATTGGAGTCTGTAGGCTTCTGTAT 43453
QY 1660 CTTGTATAGCATCTCTTCTCAAGGTAGGAAATTTTCTTTTGTGTTTCTTGAATAAT 1719
Db 43454 GTTCAATGGCATCTCTTCTTCTGTTTGGGAAGTTTCTTCTATATTTTGTGAAGAT 43513
QY 1720 ATTTTCCCTGCTTTTGACCTG-----CCTTCTTCCCTTCTCTATCTT 1765
Db 43514 ATTTCTGCGCCCTTTAAGTTGAAAATCTTCTATCTCATCCACTCTTATATCCATAGTT 43573
QY 1766 TGGTTTTTGCATAGTCTCTGCTTCTGCTTCTGATGTTTATCCCTGGATTTATTTAGACTT 1825
Db 43574 TGGTCTTCTCATGTTCTGTGATTTCTGATGTTTGTAGTTAGGATCTTTTGTGATTT 43633
QY 1826 AACATTTTCTTGACCAAGGTATCCATTTCTTCTATCTTGTCTTCACTGCTCGATTTCT 1885
Db 43634 TGCAATTTCTTGTATTGTGTGCTCATGTTCCCTACAGAAATCTTCTGCACCTGACATTT 43693
QY 1886 CTCCTCTATCTCTGTATCTGTGCTGAGTGTCTGCTGAGGTCTCTG-TTGGTTCT 1944
Db 43694 CTCCTTCCATCTCTGTATCTTGTGCTGATCTGATGTTTGTGATGTTTCTTCTTCTTCC 43753
QY 1945 TAAATTTTCTATCCAGATTTCTTCAATTTCTGTTGGTGTGTTTATTAATCTATTTCCAC 2004
Db 43754 TAGGTTTCTATCTCCAGTGTGCTCACTTGGTGTCTTATTTGTGCTACTTCCCT 43813
QY 2005 TTTCAAGTCTGAAATGTTTACTCATTTTCTCTCCAGTATTT---TACATTTCTATAGT 2061
Db 43814 TTTGAGTCTAGTATAGTTTGTTCATTTCCATCCACTGTTTGGATTTGCTTCTCTCTT 43873
QY 2062 TCTTTTAATGATTTATTTCTTCTTCTTCAAGGACCTTTTATGAATTCATATAATGAT 2121
Db 43874 TCTTTAAGGACTTCTACTCTGTTGATTTGTTTCTTCTGTTTCTTTTAAAGGACTTGTAA 43933
QY 2122 GTTAAG 2127
Db 43934 GTCAAG 43939

RESULT 41
US-10-175-523-62/c
; Sequence 62, Application US/10175523
; Publication No. US20030096264A1
; GENERAL INFORMATION:
; APPLICANT: Brockman, Jeffrey
; APPLICANT: Evans, David
; APPLICANT: Hook, Derek
; APPLICANT: Klimczak, Leszek
; APPLICANT: Laeng, Pascal
; APPLICANT: Palfreyman, Michael
; APPLICANT: Rajan, Prithi
; TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)
; FILE REFERENCE: 3235/IJ795-US3

; CURRENT APPLICATION NUMBER: US/10/175,523
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 60/299,151
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/317,828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/325,150
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/333,047
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 60/349,936
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/361,834
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62
; LENGTH: 185548
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-175-523-62

Query Match 24.0%; Score 546,799; DB 1; Length 185548;
Best Local Similarity 62.2%; Pred. No. 1.9;
Matches 1156; Conservative 0; Mismatches 572; Indels 132; Gaps 20;

QY 292 TCATTTCTTGTATTTCTATCTTTGGCTCATTTTAACTCAGTAGTGTGTTGGTTTCC 351
Db 129820 TTTCTTTCTTTATTTCTTCTTGAACCAAGTGTCAATTAAGAGAGTGTGTTCAGTTTCC 129761
QY 352 ATAAGTTTCTAGTTTCTGTTTCTGTTCTGTTGTTGTTGTTTATCTAGATTTAAGCTGT 411
Db 129760 ACCTGAATGTGGCTTTCCATTTATTATG-----TGTATTGAAGATCACCTTAGTCCAT 129705
QY 412 GGTGGTCAGATAGGACATAGAGTATTATTTCATTTGTCTTTTATCTGCGAGACTTGCTT 471
Db 129704 GGTGGTCTCATAGGATGATGGACAAATTTCAATTTTGTATCTGTAGGCTGT 129645
QY 472 TGTTTTGAATATGTATTCATTTTGGAGA--GTTTCATAGGTCCTGCACAGAGGTAC 529
Db 129644 TGTGACCAATATATGTTCAACTTTGGAGAGGTCCCATGAGGTGCTGAGAAGAGGTAT 129585
QY 530 AGTC-TTTGTGTTTGGTGAATAGTCTGTAATATCT-CTAGGTCCACTTGGTTTATGA 587
Db 129584 ATCCTTTGTTTTAGATTAATATGTTCTGTAGATATCTGTTAAGTCCATTTGTTTCAAA 129525
QY 588 CATCAGTTAGTCCAGCATTTCTGTTTCTGTTTCTGTTTGTGTAGATGACCTAACTGTGA 647
Db 129524 CTTCTGTTAGTTCATCTGTGCTCCCTGTTT-GGTTTCTGTTTCCATGATCTGTCCATTGAT 129466
QY 648 GAGAATGGGTATGGAAGTAGCCCACTATCTGTGTGTGAGT-CAATATGATTTTATAGC 706
Db 129465 GAAAGTGTGTGTGAAGTCTCCCACTATATTGTGTAGGTGCAATGTGTGCTTTGAGC 129406
QY 707 TGTAGCTGTGCTTGTTTTATGAACCTTGGGTGACATGTGTTTGGTGCATAGACATTT-AAG 765
Db 129405 TTTACTAATGTCTTTTAAATGAATGTGGTCCCTTGCATTTTGGAGCATAGTATTTCAGG 129346
QY 766 AATGCAATGTCTCTTGGTGAATTTT-CTTTTGTGCTATGTAGTATTCTTCCCAATC 824
Db 129345 AATTGAGATTTCTCTTTGAAGGATTTTACCTTTTATGAATATGAAGTGTCCCTCTCTGTC 129286
QY 825 TCATCTGCTTAGTTTGGGTTTAAAGTCTA-TTAGTCAGATATTAAATGACITGTTACGGC 883
Db 129285 TTTTGTATAACTTTTGGGTGGAAGTCGATTTTATTCGATATTAGAATGGCTTACTCCAGC 129226
QY 884 TTGCTTCTTAGGGCAATTTGCTTAGAATA-TCTTTTCCATCCTTTTACTTAAGGTGATG 942
Db 129225 TTGTTTCTTTCAGACCTTTTGTGTTGAAATTTGTTTCCAGCCTTTTCACTCTGAGGTAGTG 129166
QY 943 TCT--ATCCATGGTAGGTGTCTTTTGGATGCGACAGTAGGATCTTGTGTTTCTAT 1000
Db 129165 TCTTCTTTTCCCTGAGATGGGTTTCCCTGTAAGCAGCAATGTTGGTTTCTGTTTGTGT 129106

Db	51129	AAAGTTTTTAAATTCCTTATATTCCTCAATGACGAAGTTATCAATCAGTAGACGGTGTGTC	51188
Qy	345	GGTTTCCATAAGATTTGTAAGATTTTCTGTGTTTCTGTTCGTGTCGTGTTATCTAGATTT	404
Db	51189	AGCTTCCATGTGATGTGCTTCCCTTGTT- - -TTCTTGGTATTGAAGACCGGCTTT	51244
Qy	405	AAGCTGTGTGTCAGATAGGACATAGATATATTTCAATTGTCCTTTTATCTGTCTGAGA	464
Db	51245	ACTCCATGGTGATCTCATAGGATGCAATGGATATATTTCAATCTTCTGTATCTGTGTAGG	51304
Qy	465	CTTGCTTTGTTTGAATAATGATTTCAATTTTGGAGA- -GTTTCAATAGGTCGTGACAAG	522
Db	51305	CCGTGTTTGTAAACCAATAATATGGTCAGTTTTCGAGAAGTACCATGAGTGTCTGCGAG	51364
Qy	523	AAGGTACAGCTTTGTGTTTTT- -GGTGAATAGTCGTGTAATA- -TCTCTAGGTCCACTT	578
Db	51365	AAGGTACATTCATTTTGCCTTAGGATGAATGTCTATAAATAATCTGTAAACGCTTTA	51424
Qy	579	GGTTTATGACATCAGTTAGCTCCAGCATTTCTCTGTTTCGTTTTGTTTGTGATGACCTA	638
Db	51425	GGTCTTAACTTCTATTAGTTTCACTGTGTCCTGTTT- -TTCTGTTTTCATATCTG	51483
Qy	639	ACTGTTGGAGAGAAATGGGTATTAAGTAGAGCCCATCTCTGTGTGTGAGGT- -CAATATGT	697
Db	51484	TCCATTTGCTGAGAGTGGGGTGTGAAGTCTCCTACTATTTATTTGTGGGGTCAATGTGT	51543
Qy	698	GATTTTAGCTGATGCTGCTGTTTTTANGAATTGGGTGACATTTGTTTGGTGCATAG	757
Db	51544	GCCTTGAGCT- - -TAAGTCTCTTATGCAATGTGTTTCCCTTGCATTTGGAGCACAG	51598
Qy	758	ACATTAAGAATTCGAATGCCTCTGGTGG- -TTTTCCTTTGATGCCATGTAGTATTTCT	816
Db	51599	ATCTTCAGAATTCAGAGTTTCATTTTGGTAGATTTTCCCTTTCAGTAGTATGAAGTCTCT	51658
Qy	817	TCCCAATCTCATCTGCTTAGTTTGGGTTTAG- -TCTATTAGTCAGATTAATAATGACT	875
Db	51659	TCCTTATCTTTTGTGATACTTTTGGTGGAAAGTGTGTTTTATATAGACTAGAAATGGCT	51718
Qy	876	GTATCGGCTGCTTCTTAGGGCCATTTGCTTAGAATA- -TCTTTTCCATCCTTTTACTCTA	934
Db	51719	ACTCCAGCTGTGTTCTTGGGACCAATTTGCTGAAACACTGTGTGCCATGTTTACTCTG	51778
Qy	935	AGGTGATGCTATCCATG- - -GTAGTGTGCTTTTTTGGATGCACGACGATGGATCT	991
Db	51779	AGGTAGTGTGTCGCTTTGTCACTGAGTGCCATTTCCCTATATGAACAAAGTCTGGATCC	51838
Qy	992	TGTTTTTCATATCCATCTGTATCCAGTATCTTTTTCTAGAGAAATTAAGATCATTTAGT	1051
Db	51839	AGTTTACATATCAGTTGGTTAGTCTATGCTCTTTTGTGGGGNATGA- - - -GCC	51889
Qy	1052	CATTGATGTTGAAATATCAATAGCAGTGTGTTGGATTTCTGTTATCTTGCACTGT	1111
Db	51890	CATTGATATTAAGAGATATTAGAGAAAGTAATGTTACTGCTCTTATTTTGTGTTA	51949
Qy	1112	GAAGT	1171
Db	51950	GAGGTGAATATGTTTGTGTGCTTCTCTTTTGGGTTTG- - - - -	51991
Qy	1172	TGTGTGCTGTGTGTGTCTTGTTGTGTGTGTGTGTCTCTCCCTCTTTTGTGATTTTGGCC	1231
Db	51992	- - - - -	51991
Qy	1232	TGGAATTTATTTATTTCTATATTTTCTTGAATGTGGTTAACATCTTTAGATTGAATTTT	1291
Db	51992	TTGAAAGATTTACTTTCTGCTTTTTCTAAAGTGTAGTTTCCCTCTCTGTGTGGTGT	52051
Qy	1292	TCCTCTAGCCTCTTTTAGTCTGCAATTTGAAGATAGATATCTTTTACATCTGATTTTATC	1351
Db	52052	ACAAC- - - -TTTGTAGGCTGGGTTTATGGAAGAGATTTGTGTAAATTTGGTTTGTCT	52105
Qy	1352	TTAGATATGCTTTCTTCTCCACATATTGTGTGACAGAAAGTTTTTCTAAGTCGATGTCT	1411
Db	52106	ATGGAATATCTGGTTTCTCCATCTATGTGTAATTGAGAGTTGTGCTGGGTATAGTAGCT	52165

RESULT 44

US-10-374-077-209
: Sequence 209, Application US/10374077
: Publication No. US20040006779A1
: GENERAL INFORMATION:
: APPLICANT: Fu, Ying-Hui
: Yu, Chang-En
: Oshima, Junko
: Mulligan, John T.
: Schellenberg, Gerald D.
: TITLE OF INVENTION: ANTIBODIES AGAINST GENE PRODUCTS RELATED TO
: WERNER'S SYNDROME
: NUMBER OF SEQUENCES: 209
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Seed Intellectual Property Law Group
: STREET: 701 Fifth Avenue, Suite 6300
: CITY: Seattle
: STATE: Washington
: COUNTRY: USA
: ZIP: 98104-7092
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible

QY 2059 GCTTCTCTTAATGATTTATTCATTTCTCTTCAAGGACCTTT 2101
Db 27844 TAAATCTGTAAGGAATTTTGTGTTCTCTCTTTAAGGCTTCT 27886

RESULT 45
US-10-085-117-61/c
; Sequence 61, Application US/10085117
; Publication No. US2003023234A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000121
; CURRENT APPLICATION NUMBER: US/10/085,117
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 96602
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: variation
; LOCATION: (1)...(96602)
; OTHER INFORMATION: n = any nucleotide
US-10-085-117-61

Query Match 22.9%; Score 521.599; DB 1; Length 96602;
Best Local Similarity 61.8%; Pred. No. 3.3;
Matches 1157; Conservative 0; Mismatches 579; Indels 136; Gaps 22;
QY 289 ACTTCTATTTCTGATTTCTATCTGCTCAATTTTAACTCAGTGTGAGTGTGTTGTT 348
Db 5698 AATTTCTTTCTTTATTTCTTCCCTTGACCAAGTTATCATTCAGTAGAGTGTGTTGCT 5639
QY 349 TCCATAAGTTGTAAGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTAA 408
Db 5638 TCCATGTAAGTGTGCTTCCATGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTT 5582
QY 409 TGTGCTGTGATAGGACATAGAGTATTTTCAATTTGCTTTTATCTGTCGAGACTTG 468
Db 5581 AGTGGTGTGATAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5522
QY 469 CTTTGTGTTGAAATGATTTCAATTTTGGAGA--GTTTCATAGGGTGTGCTGACAGAGG 526
Db 5521 TTTTGTGACCAATATATATGTTGCTCAATTTTGGAGAAGGTACAGTGTGCTGCTGAGAAG 5462
QY 527 TACAGTCTTTGCTGTTTGG--TGAATAGTCTGTAATATC--TCTAGTCTCACTTGGTTA 584
Db 5461 TATATCTTTTCTTTAGGATGAATGTTCTATATAATCAATGATGATCTATTTGTTCA 5402
QY 585 TGACATCAGTGTGCTCAGCAATTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTT 644
Db 5401 TAACTTCTGTTAGTTTCACTGTGCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTT 5343
QY 645 GGAGAGATGGGTATGAACTAGCCACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 704
Db 5342 GCTAAGAGTGGATGTTAAAGTCTCCACATATATGAGTGGGT--GCCATGTTGCTAA 5284
QY 705 GCTGTAGCTGTGCTGTTTATGAACCTGGGTGACATTTGTTTGGTGCATAGACATTA 764
Db 5283 GCTTTAGTAAAGTTCTCTTATGAATGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5224
QY 765 GAATTCGAATGCTCTTGTGGGA--TTTTCCTTTGATGCTGCTGCTGCTGCTGCTGCTGCT 823
Db 5223 GAATTTGAGGTTTATTTGTTGTTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTT 5164
QY 824 CTCATCTGCTAGTTTGGGTTAA--GTCATTTAGTACATATTAATGATGCTGCTGCTG 881
Db 5163 CATTTTGTAACTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 5104

QY 882 GCTTCTTCTTAGGCGCATTTGCTTAGAATA--TCTTTTCCATCCTTTTACTCTAAGGTGA 940
Db 5103 GCTTGTGTTCTTGGACCGTTTGTCTTGGAAAATTTGTTTCCATCTTTTACTCTGAGGTAG 5044
QY 941 TGTCTATCCAAG--GTAAGTGTCTTTTGTGATCAGCAGTATGATGATGATGATGATGAT 997
Db 5043 TGTCTGTCTTTGTCACTGAGATGCACTTCTGTATGAGCAAAATGCTGGTCTCTGTTTA 4984
QY 998 CATATCCATTTCTGTTACCCAGTATCTTTTCTAGAGAAATTAAGATCATTAGT--CATGG 1056
Db 4983 TGTATCCAGTCCGTTTATCTATGCTCTT-----TTTGGATAATTTAGTCCATTG 4935
QY 1057 ATGTTGAGAAATATCAATGAGCAGTGTGTTGTGATTTCTTGTATCTTGCACCTTGTGAAGT 1116
Db 4934 ATGTTAAGAGATATTAAAGAAAAGTGAATTTACTTCTGTTTATTTTGTGTTAGAGGC 4875
QY 1117 GT 1176
Db 4874 AGAATTATGTTGT 4838
QY 1177 GTCTGT 1236
Db 4837 -----TTGAA 4833
QY 1237 TTTATTTATTTCAATTTTCTTGAATGGGTAAACATCTTTAGATGAGTTTCTCTCC 1296
Db 4832 AGATTACTTTCTTGTCTTTTCTAGGGTAAATTTCCCTTCTTGTGTGTGTGTGTGTGTGTGT 4773
QY 1297 TAGCCTTCTT--TAGTCTGTGATTTGAAGATAGATATCTTTACATCTCATTTTACTTTA 1354
Db 4772 TATATCTTTTGTAGGGTGGATTTATGGAAGATATGTTGTAATTTAGTTTGTGTGTGTGT 4713
QY 1355 GAAGTCTTTCTTCTCCAACTATTGTGACAGAAAGTTTTTCTAAAGTCAGTAGTCTGGC 1414
Db 4712 GAATATCTTGTGTTTCTCCATCTATGTAATTTAGAGTGTGTTGTGGGTATAATAGACTGG 4653
QY 1415 CTGACATCTGTAGTCTCTTGGAGTCTGTAGCAGTCTGTGAGGGCTTCTTACATTGTG 1474
Db 4652 CTAGCAATTTGTGTCTCTTAGGGTGTATGACATC--ATCCAAAATCTTCTAGCTTTCTAT 4594
QY 1475 AGTTTCTATTGGAAAAGTCAGGTGTAATTTCTAATACATCTGCTTTTATATGTTAATTTGT 1534
Db 4593 AGACTCTGTG--GAGAAGTCTGGGTAGTCTCATAGGTTTGGCTTCTATGATTTACTTGAC 4535
QY 1535 CTTTCTTCCCTGTCATCTTTTAAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1594
Db 4534 CTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 4476
QY 1595 ATTATGCACTGTGGG---GAGTCTTCTTCCGTCCTCAATCTATTTGGTGTGTTGATGC 1650
Db 4475 ATTTAGACAAGGGAATGAATTTCTTCTGTCCTGCTATTTGGAGTCTCTGTAGGC 4416
QY 1651 TTTCTGTACCTGTATAGGCACTCTTTCTCAAGTTAGGAAATTTTCTTCTTCTTCTTCTTCT 1710
Db 4415 TTTCTGTATGTTTATAGGCACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4356
QY 1711 CTTGAAAATATTTTCCCTGCTTTTTCACCTG-----CCTTCTTCTTCTTCTTCTTCTTCT 1756
Db 4355 GTTGAAGATATTTCTGCGCTTTTGTAGTGAAGATCTTGTCTCTCTATACCTATTAT 4296
QY 1757 CTATTCCTTTGTTTGTGATAGTCTCTGCTGCTTCTTCTGCTGCTTCTTCTGCTGCTGCT 1816
Db 4295 CTTAGGATTCATCTTCTCATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4236
QY 1817 TTTAGACTTAACTTTCTTTGACCAAGATATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1876
Db 4235 TTTGTATTTGAATTTTGTGTTGACTGTTGTGCTCAATGTTTCTTCTTCTTCTTCTTCTT 4176
QY 1877 TGAGATTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1935
Db 4175 TGAGATTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4116

us10664775-4.rnpb

Wed Aug 25 15:11:59 2004

Qy	1936	TTGGGTTCTTAATTTTTCATTTCCAGATTCCTTCAGTTTGGTTTGTGTTTATTAATTC	1995
Db	4115	CCTCTTTCCTAGGTTTCTAACTCTACGGTTGTCTCCCTTTG---TACTTTCATTAATTC	4059
Qy	1996	TATTTCCACTTTCAGGTCCTGAAATGTTTACTCATTTTCTCC---AGTATTTACAT	2052
Db	4058	TATTTCCATTTTGAATCTTGGATGTTTGTTCATTTCCATCACCTGTTGATGTT	3999
Qy	2053	TTCATAGGTTTCTTAAATGGAATTTATTCATTTCTCTTCAAGGACCTTTTATGAATCAT	2112
Db	3998	TTCCCTAAATCTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT	3939
Qy	2113	AAATGTATGTT	2124
Db	3938	TTGTTTGTGTTT	3927

Search completed: August 25, 2004, 09:28:53
Job time : 4466 secs

Result No.	Score	Query		DB	ID	Description
		Match	%			
C	1 558.599	24.5	3843	1	AK041981	ACCESSION:AK041981
	2 556.099	24.4	3987	1	AK067425	ACCESSION:AK067425
	3 550.401	24.2	2841	1	AK046480	ACCESSION:AK046480
	4 547.599	24.0	3235	1	AK033883	ACCESSION:AK033883
	5 544.798	23.9	3025	1	AK037785	ACCESSION:AK037785
C	6 541.399	23.9	4895	1	AK088507	ACCESSION:AK088507
	7 534.359	23.3	3204	1	AK085977	ACCESSION:AK085977
	8 521.598	22.9	1644	1	AK008913	ACCESSION:AK008913
	9 515.198	22.7	2436	1	AK089845	ACCESSION:AK089845
	10 515.998	22.6	4832	1	AK046811	ACCESSION:AK046811
C	11 503.498	22.1	3375	1	AK045866	ACCESSION:AK045866
	12 481.397	21.1	4466	1	AK034168	ACCESSION:AK034168
	13 447.898	20.1	2108	1	AK041956	ACCESSION:AK041956
	14 448.597	19.7	3179	1	AK048989	ACCESSION:AK048989
	15 430.098	18.9	4502	1	AK079106	ACCESSION:AK079106
C	16 428.398	18.8	2378	1	BC021486	ACCESSION:HC021486
	17 418.399	18.4	3279	1	AK040345	ACCESSION:AK040345
	18 412.599	18.1	2535	1	AK050942	ACCESSION:AK050942
	19 399.299	17.5	1151	1	BZ124878	ACCESSION:BZ124878
	20 393.099	17.2	1065	1	BZ135754	ACCESSION:BZ135754
C	21 385.199	16.9	2372	1	AK040990	ACCESSION:AK040990
	22 363.399	15.9	886	1	BZ109340	ACCESSION:BZ109340
	23 361.399	15.9	846	1	BH348578	ACCESSION:BH348578
	24 358.399	15.7	825	1	BZ2221947	ACCESSION:BZ2221947
	25 355.899	15.6	900	1	BH308381	ACCESSION:BH308381
C	26 354.899	15.6	1037	1	BZ132018	ACCESSION:BZ132018
	27 353.899	15.5	914	1	BZ106508	ACCESSION:BZ106508
	28 351.899	15.4	1075	1	BZ135244	ACCESSION:BZ135244
	29 351.499	15.4	920	1	BZ108972	ACCESSION:BZ108972
	30 351.099	15.4	823	1	BZ109050	ACCESSION:BZ109050
C	31 349.899	15.4	816	1	BZ184947	ACCESSION:BZ184947
	32 349.699	15.3	910	1	BZ097017	ACCESSION:BZ097017
	33 349.499	15.3	061	1	BZ244766	ACCESSION:BZ244766

MEDLINE	2032030913	
PUBMED	11076861	
REFERENCE	4	
AUTHORS	<p>THE RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.</p> <p>Functional annotation of a full-length mouse cDNA collection</p> <p>Nature 409, 685-690 (2001)</p>	
TITLE	5	
JOURNAL	<p>THE FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.</p> <p>Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs</p> <p>Nature 420, 563-573 (2002)</p>	
REFERENCE	6	
AUTHORS	<p>(bases 1 to 3887)</p> <p>Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carrinci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitho,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shingawa,A., Shiraki,F., Sogabe,Y., Tagami,M., Takagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.</p>	
TITLE	<p>Direct Submission</p>	
JOURNAL	<p>Submitted (16-Apr-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222.</p>	

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.
URL:<http://genome.gsc.riken.go.jp/>

```

FEATURES
  source
    URL:http://fantom.gsc.riken.go.jp/.
    Location/Qualifiers
      1..3887
        /organism="Mus musculus"
        /mol type="mRNA"

```

```

strain="C57BL/6J"
db_xref="FANTOM_DB:E130020L12"
db_xref="MGI:243267"
db_xref="taxon:10090"
clone="E130020L12"
clone.lib="RIKEN full-length enriched mouse cDNA library"
dev stage="0 day neonate"
tissue type="eyeball"

```

```

misc_feature
1..3887-
/note="unknown EST (GB|BB797679, evidence: BLASTN, 96%,
match=243)"

Query Match      24.4%; Score 556.099; DB 1: Length 3887;
Best Local Similarity 62.0%; Pred. No. 0.7;
Matches 1144; Conservative 0; Mismatches 574; Indels 127; Gaps 18;

```

292	TCTATTCTCTGATTCTCTATCTTGGCTCATTTTAACTCAGTAGTAGTGTGTGTGGTTTCC	351
QY		
720	TTTCTTTCTTTATTCTCTCCCTTGACCAAGGTATCTTGAAGAAGAGTGTGTTTCAGTTTCC	779
DB		
352	ATAAGTTGTAGATTTTCTGTGTGTTTCTGTGCTGTGTGTGTATCTAGATTTTAAAGCTGT	411
QY		
780	ACGTGAATGTTGGCTTCTTATTATTATGTTGTTATTGAAGAT--CAACCTTAGTTCAT	836
DB		
412	GGTGTCTGATAGGACATAGAGTATTATTTCAAATTGTCTTTTATCTCTCGAGACTTGCCT	471
QY		
837	GGTGATCTGATAGGATGATGGGCAATTTTCAATATTTTGTGCTGTGTGTGAGACTGTGTT	896
DB		


```

QY 1409 TCTGGCCTGACATCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGCGAGGCTCTCTTAC 1468
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1214 CTTGGGTTGGCATTTGTCTCCCTTAGGCTCGCATACATCTGTGCCAGGATCTCTTGGC 1155
QY 1469 ATTTCAGTTCCTATTGGAAGAGTCAGGTGAATCTTAATCATCATCTGCTTTATATGTTA 1528
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1154 TTTCATAGTCTCTGTG-GAAAAATCTGTGTAAATCTCATAGGCTGCTTTATATGTTA 1096
QY 1529 ATTGTCCTTTTCCCTTGCATCTTTTAATATCTTTCTTTGTTCTATACCTTTTAGTGAAT 1588
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1095 CTT-GACCTGTTCCCTTACTCTCTTTAATATCTTATCTTATTTAGTGCAATTTTGTGTT 1037
QY 1589 TTGATTATATGACACTGTGGGAGTCTTTTCCGGTCCAACTATTTTGGTGTGTTTGTAT 1648
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1036 CTGATTATATGTCGGGAGGAATTTCTTTCTTGTCCAGTCTATTTGGAGTCTCTGTAG 977
QY 1649 GCTTCTTGTACCTTCATAGGATCTCTTCTCAAGGTTAGAAATTTTCTTTTGTGTT 1708
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 976 GCTTCTTGTATGTCATGSCATGTCATCTTTAGGTTTGGAGTCTCTTCTTATATTT 917
QY 1709 TCTTGAATAATTTTCCCTGCTTTTGACCTG-----CCTTCTCCCTTC 1754
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 916 TTGTTGAAGATATTTGCTGTCTTAAAGTTGAAATTTTCAATCTCATTTTCCCTTATTT 857
QY 1755 CTCTATCTCTTGGTTTGTGATAGTCTCTGCTTCCCTGATGTTTATGCTCTGATTT 1814
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 856 GTCTGTAGTTTGGTCTCTCATCTGCTCTGATTTCTCTGATGTTTGTAGTTAGATC 797
QY 1815 ATTTPAGACTTAACATTTCTTTTGACCAAGGATTCATTTCTTCTATCTCTCTCACTG 1874
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 796 TTTTGTGATTTTGCATTTCTTGTATGTTGTCCTGATTTCTCTGATGTTTGTAGTTTCTGCA 737
QY 1875 COTGAGATCTCTCTCTATCTCTGTATCTGTCTGATGAGGCTGTCTCTGAGGTTCTT 1934
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 736 COTGAGATCTCTCTCTATCTCTGTATCTGTCTGATGAGGCTGTCTCTGATGTTTGTAGTTTCTGCA 677
QY 1935 G-TTGGGTTCTTAATTTTTCATTTTCCAGATTTCTTCCAGTTTGGGTTTGTGTTTATTAAT 1993
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 676 GATTTCTTCTTAGGTTTATATCTCCAGCATTTGCTCACATTTGGGT---TTTATTTGTT 621
QY 1994 TCTATTTCCACTTCAGTCTCGGAATGTTTATCTATCTTCTCTCCCTCC---AGTATTTACA 2050
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 620 TCTACTTCCCTTTTAGTCTCTTGTATGTTTGTTCAGTCCATCACCCTGTGTTGATTTG 561
QY 2051 TTTTTCATAGTTTCTTAAATGATTTATCTATTTCTCTTCCAGGACCTTTTATGAATTC 2110
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 560 TTTTCTCTATTTCTTTAAGACCTTCACTCTTATCATGATGTTCTGTATTTCTTTA 501
QY 2111 ATAAATGATGTTAAGTCTCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTCA 2163
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 500 AGTGAGTTATTAAAGCCCTTCTTGATGTCCTCTTCCAGCATCATGAGATATGA 448

```

RESULT 4

```

LOCUS AK033883
DEFINITION Mus musculus adult male diencephalon cDNA, RIKEN full-length
          enriched library, clone:9330109E13 product:unclassified, full
          insert sequence.
ACCESSION AK033883
VERSION AK033883.1 GI:26083632
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636

```

REFERENCE

AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159

REFERENCE

AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuda, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multi-capillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861

REFERENCE

AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409, 685-690 (2001)

REFERENCE

AUTHORS The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 JOURNAL Nature 420, 563-573 (2002)
 MEDLINE 12033335

REFERENCE

AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Inctani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
 TITLE Direct Submission

REFERENCE

JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@sc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site for further details.
 URL:http://genome.gsc.riken.go.jp/
 URL:http://fantom.gsc.riken.go.jp/

FEATURES

source
 1..3235
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM DB:9330109E13"
 /db_xref="MGI:2398061"
 /db_xref="taxon:10090"
 /clone="9330109E13"
 /sex="male"
 /tissue_type="diencephalon"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"

misc_feature	/dev_stage="adult"	1..3235	/note="unclassifiable"
Query Match	24.0%;	Score 547.599;	DB 1; Length 3235;
Best Local Similarity	62.2%;	Pred No. 0.9;	Mismatches 594; Indels 100; Gaps 19;
Matches 1144;	Conservative	0;	
QY	292	TCTATTTCTGATTTCTATCTATCTGGCTCAATTTTTAACTCAGTAGTGGATTTGTTGGTTTCC	351
Db	66	TTTCTTTCTTTATTCCTTCTTGACCAAGTTATCATTCAGAAAAAGTTTGTTCAGATTCC	125
QY	352	ATAAGTTTGTAGTTTCTGTTGTTTCTGTTGTTGTTGTTGTTGTTTATCTATAGATTTAAGCTGT	411
Db	126	AGTGAATGTGGCTTTCCGTTATTTATTTGTTTTTTGAAATC---AGCCTTAGTCCAT	182
QY	412	GGTGTCAGATAGGACATAGAGTATTTATTTCAATTTGCTTTTATCTCTCGAGACTTGCTTT	471
Db	183	GGTGATCTGATAGGATGATGGGAAAATATCAATATTTTGTATCTGTGAGTCCCTGTTT	242
QY	472	TGTTTTGAATATGATTCAAATTTTGGAGA--GTTTCATAGGGTCTGACAGAGGTAC	529
Db	243	TGTGACCAATATATGGTCAGTTTGGAGAGGTCCTCAGAGGCTCAGAGCAGGAT	302
QY	530	AGTC--TTTGTGTTTTGGTGAATAGTCTGTAATAATCT--CTAGTCCACTTGGTTTATGA	587
Db	303	ATCCTTTTGTTTAGGATAAATGTTCTATAGATATCTGTTAGATTCATTGTGTTTATAA	362
QY	588	CATCAGTTAGCTCCAGCATTTCTCTGTTTCGTTTTTGTGAGATGACCTAACCTGTTGGA	647
Db	363	TGTCCTGTTAGTTTAAATGCGTCCCTGTTTGAG--TTTCTGTCCTCCATGATCTGTCCATTGGT	421
QY	648	GAGAATGGGATATGAAGTAGCCACTATCTGTGTGAGGT--CAATATGTGATTTTTPAGC	706
Db	422	GCAAGTGTGTGTTGAGTCTCCCAATATTTATGTGAGGTACAAATGATGCTTTTGAGC	481
QY	707	TGTAGCTGTGCTGTTTTATGAACCTGGGTGACATTTGTTTGTGTGCATAGACATTAAGA	766
Db	482	TTTACTAAAGTTTCTTTAAATTAATGTGCTGCCCTTGTATTTAGAGCATAGATATTACGA	541
QY	767	ATTGCAATGCTCTCTGTGGGATTTT--CCTTTGATGCTATGTAGTATTTCTTCCCATCT	825
Db	542	ATTGATAGTTCTCTTTGAGGATTTTACCTTTGATGAGTATGAAGTCCCTCCCTTATCT	601
QY	826	CATCTGCTTAGTTTTGGGTTTAAAGTCTA--TTAGTCAGATATTTAAATGACTGTATCGGCT	884
Db	602	TTTTTGTACACTTGGTGTGGAAGTCGATTTTATTAGATATTAGAAATGGCTACTCCAGCT	661
QY	885	TGCTTCTTAGGCCCATTTTGCTTAGAATA--TCCTTTCAATCCTTTTACTCTAAGGTGATGT	943
Db	662	TGTTTCTTCGGAGCATTTGCTTTGAAAAATGGTTTCCAGCCTTTTATTTTGGAGTAGTGT	721
QY	944	CTATCCAT---GOTAGTTGTCTTTTTTGGATGAGAGTAGGATGGATCTTGTTTTTCAT	1000
Db	722	CTATCTTTTCTCTGAGATGATGTTTCCAGTAAGCAGCAGATGTTGGGCTCTGTGTGT--GT	780
QY	1001	ATCCATTTGTTTACCAGTATCTTTTTCTAGAGAAATTAAGATCAATGAGTCAATGATGT	1060
Db	781	AGCAGTCTGTAGTCTATGTCTTTTTTATTTGGGAGTTGAG-----TCCATTGATAT	832
QY	1061	TGAGAAATATCAATAGAGAGTGTGTTGATTTCTTGTTATCTTGCACCTTGTGAAGTGTGT	1120
Db	833	TAAGAGATATTAAGGAAAGTAATTTGTTGCCCTCTTATTTT-----	875
QY	1121	GT	1180
Db	876	-----TGTGCTAAAGTTGGCATTTCTCTTCTTGTGGCTGTCTCTTTTAGGTTT	924
QY	1181	GTGCTGT	1240
Db	925	GTTGAGGATTAACCTTCTCTCTTTTTCAGGACTGTTTTTCAGTCCCTTGAATGTTTTTT	984
QY	1241	TTATTTATTCATATTTTCTTGAATGGGTAAACATCTTTAGATATGAAGTTTTTCTCCPAGC	1300

[illegible][illegible]

```

FEATURES
  source
    Location/Qualifiers
      1. 4895
        /organism="Mus musculus"
        /mol_type="mRNA"
        /strain="NOD"
        /db_xref="FANTOM_DB:E430019C06"
        /db_xref="MGI:2427404"
        /db_xref="taxon:10090"
        /clone="E430019C06"
        /cell_type="thymic cells"
        /tissue_type="thymus"
        /clone_lib="RIKEN full-length enriched mouse cDNA library"
        /dev_stage="2 days neonate"
      1. 4895
        /note="unclassifiable"

  misc_feature
    Query Match      23.9%; Score 544.399; DB 1; Length 4895;
    Best Local Similarity 61.7%; Pred. No. 0.63;
    Matches 1148; Conservative 0; Mismatches 576; Indels 136; Gaps 19;

QY 292 TCTATTTCTTGATTTCTATCTTGCTCATTTTAACTCAGTAGTGTGTTGGTTTC 351
DB 932 TTTCTTTCTTTATTCCTTCCTTGACCAAGGTATCATTCAGAAAGAG-----CC 978
QY 352 ATAAAGTTGTAAGTTTCTGTTGTTTCTGTTTGTGTTGTTGTTATCTAGATTAAAGCTGT 411
DB 979 ACGTGAATGTTGGCTTCTATATTTATTTATTTGTTTATTGAAGATCT---GCCTTAGTCCAT 1035
QY 412 GGTGTCAGATAGACATAGAGTATTTTCAATTTGCTTTTATCTCTCGAGACTTTCCTT 471
DB 1036 GGTGATCTGATAGATGATGATGACAAATTCATATTTTGTATCTATTGAGGCCCTGTT 1095
QY 472 TGTGTTGAAATATGATTCATTTTGGAGA--GTTTCATAGGGTGTGTCAGAAAGGTAC 529
DB 1096 TGTGACCAATATATGTTCAATTTTGGAGAAAGGTACCATGAGGTGCTGAGAAAGGTAT 1155
QY 530 AGTC--TTTGTGTTTGGTGAATAGTCTGTAATATCT-CTAGTCCACTTGGTTTATGA 587
DB 1156 ATCCTTTGTGTTTGGTGAATAGTCTGTAATATCT-CTAGTCCACTTGGTTTATGA 587
QY 588 CATCAGTTAGTCCAGCATTTCTCTGTTTCGTTTGTGTTGTTGATGACACCTAACTGTTGGA 647
DB 1216 CTTCTGTTAGTTTCACTGTGCTCCTGTTTAG--TTCTGTTTCTGCGATCTGTCCATTGAT 1274
QY 648 GAGAAAGGATATGAAATGACCCACATCTGCTGTGAGGT-CAATATGTGATTTTATG 706
DB 1275 GAAAGTGTGTGTTGAACTCTCCACATTTTATGTTGAGGTGCAATGTGTTGCTTGA 1334
QY 707 TGTAGCTGTGTTGTTTATGAACTTGGGTGACATTTGTTTGGTGCATAGACATTAAGA 766
DB 1335 TTATTAAAGTCTTTTAAATGAATGCTGCTGCCCTTGATCTGAGCATAGATATTCAA 1394
QY 767 ATTGCAATGCTCTTGTGAGATTTCTTTGATGCCCTATGTAGTATCTTCCCAATCTC 826
DB 1395 TTGAGAAATTCCTTTTGAAGAATTTTACCTTTGATGAGTATGAAGTCTCTTCTTGTCTT 1454
QY 827 ATCTCCTTAGTTTGGGTTTAAAGTCTA--TTAGTCCAGATATTAATAAGTCTATCGGCTT 885
DB 1455 TTTTGAATAACTTTGGTTGGAAGTCGATTTTATTCGATATTAGATGGCTACTCCAGCTT 1514
QY 886 GCTCTTAGGGCCATTTGCTTAGAATA--TCCTTTCCATCTTTTACCTCTAAGGTGATGTC 944
DB 1515 GTTCTCTCAGACCAATTTGCTTGGAAATTTGTTTTCAGCCTTTTCACTCTGAGGTAGTGC 1574
QY 945 TATCCAT---GGTAGTGTCTTTTTCGATGCGACGATGAGTGGATCTTGTGTTTCATA 1001
DB 1575 TGTCTTTTCCCTGAGATGGGTTTCCCTGTAAGCAGCAATATTGGGTCTCTGTTTGTGTA 1634
QY 1002 TCCATTTCTGTTACCCAGTATCTTTTCTAGAGAAATTAAGATCATGTAGTCAATTGATTT 1061
DB 1635 GCCAGTCTGTTAGTCTATGCTCTTTTATTGGAATTTAG-----TCCATTGATATT 1686
QY 1062 GAGAAATTCAATGAGCAGTGTGTTGCGAATCTTCTTATCTTGACACTGTGTAAGTGTG 1121

```

```

DB 1687 AAGAGATATTAAAGGAAAGTAATTTGCTTCTCTATTATTTTGTGTTAGAGTTGGCAT 1746
QY 1122 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1181
DB 1747 TCTGTTCTTGTGGCTGTCTCTTTTAGGTTG----- 1778
QY 1182 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1241
DB 1779 -----TTGAAGGATTT 1788
QY 1242 TATTATTCATATTTCTTGAATGTGGTAAACATCTTTAGATTGAAGTTT---TCTCCTA 1298
DB 1789 ACTTCTCTGCTTTTCTAGGACGTGTTCCGCTCTTATTTGTTTTCCTGTTTATTT 1848
QY 1299 GCTTTCTTTAGGTCTGCATTTTGAAGATAGATATTTTACATCTCATTTTATCTTAGAAT 1359
DB 1849 ATCCTTTGAAGGGCTGGATTCATGGAAGATAATGTGTAATTTGGTTTGTGACATAGAAT 1908
QY 1359 GTCCTTCTTTCTCCAACTATTCTGACAGAAAGTCTTTCTAGTGCAGTAGTCTGSCCTGA 1418
DB 1909 ACTTTGGTTTCTCCGCTCTATTGTAATTTGAAAGTCGGGCTTGTATAGTAGCCTGGCCG 1968
QY 1419 CATCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGAGGGCCTTTCTTACATTTTACGTT 1478
DB 1969 CATTTGTGATCTCTTAGTGTCTGTATACATCTGTCCAGACTCTTCTGCTTTTATAGTA 2028
QY 1479 TCTATTGGAAGAGTCAGGTGTAATTTCTAATACATCTGCTTTTATATGTAATTTGTCCTTT 1538
DB 2029 TCTGTT-GAAAAATCTGTTGTAATTTCTGATAGGCCCTTCTTTTATATGTTACTT-GACCTT 2086
QY 1539 TTTCCCTCCATCTTTTAAATATCTTCTTCTTCTTCTATCTTTTATAGTGAATTTGATTATA 1598
DB 2087 TTCCCTTACTGCTTTTAAATATCTTCTTATTTAGTGAATTTGCTGTCTGATTATCA 2146
QY 1599 TGCACGTGCGGAGTTCTTTTCCGGTCCAAATCTAATTTGGTGTGTTTGTATGCTTCTTGTGA 1658
DB 2147 TGTGTGCGGAGGAATTTCTTTCTGGTCCAGTCTAATTTGGAGTCTCTGAGGCTTCTTGTGA 2206
QY 1659 CTTGTATAGGCATCTCTTCTCAAGGTAGGAAATTTTCTTTTGTGTTTCTTGTGTTTCTTGA 1718
DB 2207 TGTTCATGGGCACTCTTTCTTTGTTGGTTTGGAGTTTCTTCTAATAATTTGTTGAAGA 2266
QY 1719 TATTTTCCCTGCTTTTGACCT-GCCTTCTTCCCTCTCTATTC-----T 1764
DB 2267 TATTTGCTGGCCCTTTAAGTTGGAATCTTCACTTCTCTATACCTATTGTCTTAGGT 2326
QY 1765 TTGGTTTGTGATAGTGTCTGCTTCTGCTGATGTTTATGCTTGGCTGGATTATTTTAGACT 1824
DB 2327 TTGGTCTCTCATTTGTCTCTGATTTCTGGAATTTCTGGAATTTTGTAGTAGGAGCTTTTGTGCA 2386
QY 1825 TTAACATTTCTTTGACCAAGGTATCCATTTCTTCTATCTTGTCTTCACTGCTGAGATTC 1884
DB 2387 TTGCATTTCTTTGATTTGTTGCGCAATTTCTCTATGGAATCTTCTGCACTGAGATTC 2446
QY 1885 TCTCTTCTATCTGTATCTGTGTCAGTAGGAGCTGTCTCTGAGGTTCTGTG-TTGGGTTCT 1943
DB 2447 TCTCTTCCATTTCTGTATTTCTTTTGTGATCTGTCATCTATGTTTCCAGATTTCTTTTC 2506
QY 1944 TTAATTTTCTTCTTCCAGATTTCTTCTGATTTGGTTTGGTTTGTGTTTAAATTTCTATTCCA 2003
DB 2507 CTAGGGTTTCTATCTCCAGTGTGCTTCACTTTGGGTTTCTTTATTTGTGCTACTTCCC 2566
QY 2004 CTTTCAAGGTCCTGAAATGTTTACTCATTTTCTCTCCCTGAGTATTT---TACATTTTCTAGG 2060
DB 2567 TTTTATAGGCTAGTAGTGTGTTTGTCTTCTATCACCTGTTTGGATGTGTTTCTCTGTT 2626
QY 2061 TTTCTTAAATGATTTATTTCAATTTCTCTTCTCAGGACCTTTTATGAATTCATAAATATGA 2120
DB 2627 TTTCTTTAAGGACTCTTCTCTGTTTGTGTTGTTATTTTCCCTGTTTCTTTAAGGACTTGA 2686

```

AK085877 3204 bp mRNA linear HTC 20-SEP-2003
LOCUS Mus musculus 16 days neonate heart cDNA, RIKEN full-length enriched
DEFINITION library, clone:D830021017 product:unknown EST, full insert
sequence.
AK085877
KEYWORDS HTC; CAP trapper.
VERSION AK085877.1 GI:26103041
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
MEDLINE
PUBMED 10349636
2
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
MEDLINE
PUBMED 10349636
3
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuunai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yanamoto,K., Matsunoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuyama,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
Genome Res. 10 (11), 1757-1771 (2000)
20530913
MEDLINE
PUBMED 11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6
(bases 1 to 3204)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ono,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-rse@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
7
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-rse@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
8
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-rse@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)

Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
FEATURES
Location/Qualifiers
1..3204
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:D830021017"
/db_xref="MGI:2423132"
/db_xref="taxon:10090"
/clone="D830021017"
/tissue_type="heart"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="16 days neonate"
misc_feature
1..3204
/note="unknown EST (GB|BBS17954, evidence: BLASTN, 97%,
match=300)"
Query Match 23.3%; Score 531.598; DB 1; Length 3204;
Best Local Similarity 63.1%; Pred. No.1.1; Gaps 20;
Matches 1108; Conservative 0; Mismatches 524; Indels 124; Gaps 20;
376 TTCTGTGTGTGTGTGTGTATCTAGATTTAAGCTGTGGTGTCT--AGATAGGACATAGA 432
8 TTGTTTGTGTGTATTTTAAAGCAGCCCTAGTCTGTGGTGTCTGGTGTGTGTGCAATG 67
433 GTATTATTCAATGTCTTTTATCTGTGAGACTGTCTTTTGTGTTTGAATATGTATTCAA 492
68 GGATTATTCAATCTTCTTGTATCTGTGTGGCCCTGTTTGTGACCAATATATGTGTCAA 127
493 TTTTGGAGAGTTTCATAGGGTCTGACAGAGAGGTACAGTCTTGTGTTTGTGTAATA 552
128 TTTTGGAGAGGTACCATGAGTGTGTGAGAGGTATATCTTTTGTGTTTGTGTAATA 187
553 GTCGTAAATATCT-CTAGTCCACTTGGTTTATGACATCAGTTAGCTCCAGCATTTCTC 611
188 TTCTATAAATATCTGTGTAGATCCATTTGTTTCTATAACTT-GTTAGTTTCACTGTATCTC 245
612 TGTTCGTTTTTGTGTGAGTACCTAATCTGTGTGAGAGAGTGGGTATTGAAGTAGCCC 671
246 TGTGTAG-TTTCGTGTTCCATGATATGTCATGTGTTGTTGAAAGTGGGTGTGAAATCTCCC 304
672 ACTATCTGTGTGTAGGT-CAATATGTGATTTTGTAGTGTGTGTTTGTGTTTATGAAC 730
305 ACTATTATTGTGTGGGT 363
731 TTGGGTGACATTTGT 789
364 GTGGGTGTCTTGAATTTGGAGCATAGATGTTTCAGAAATTCAGAGTTTCATCTTGGTAGATT 423
790 TTTCCCTTTCATGCTATGATGATTTCTTCCCAATCTCATCTGCTTGTAGTTTGGGTTTAG 849
424 TTTCCCTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 483
850 TCTATTAGTCAGATTTAAATGACCTGTATCGGTTGTCTTCTTAGGGCAATTTGCTTAGA 909
484 TGATTGTATTCAATATTAGATGGCTACTCCAGCTTGTCTTGTGGACCAATTTGCTTAGA 543
910 ATA-TCTTTTCCATCTTTTACTCTAGGTGATGCTATCCATG---GTAGTTGTCTT 964
544 AAATTGTTTCCAACTTTTACTCTGAGGTAGTGTCTGACTTTGTCACTGAGGTGTGTTT 603
965 TTTTGGATGCAGCAGTAGGATGGATCTTGTGTTTTCATATCCATCTCTGTACCCAGTATCTT 1024
604 CTTGTGTATCCAGCAAAATGCTGGTCCCGTTTATGTAGCCAGCTCTGTAGTCTATCTT 663
1025 TTTCTAGAGAAATTAAGATCATTTGATGATGATGATGATGATGATGATGATGATGATG 1084
664 TTACTAGAGATTGA-----GTCCATTAAATGTTAGAGATATTAAAGAAAGTGTAT 714

Qy	988	ATCTGTTTTCATATCCATTCTGTATCCAGTAATCTTTTTCTAGAGAAATAAGATCATTT	1047
Db	1035	GTCTTGTTGTGTAGCAGTTCTTAGTCTATGTCITTTTTATTGGGAGTTGAGA-----	981
Qy	1048	GAGTCATTCAGTGTGAGAAATTAACAATGACCACTGTTTGTGGAATCTWTGTATCTTCGCAC	1107
Db	980	---CCAATTGATTTAAGAGAAATTAAGGAAAGTAATGTGCTTCCTGTTAT-----	931
Qy	1108	TTCGSAAGTG	1167
Db	930	-----TTATGTTGTAAGAGTTGGCAATCTGTTCTGTGCGCTGTCTCTTTTAGGGTTG	878
Qy	1168	TGT	1227
Db	877	GTTGAGGGNATCTCTTGTGTATTTCTAGGGCGATGTTCCCGTTCWTSANATGTGTTTTA	818
Qy	1228	GGCCTGGAATTAATTAATTAATTCATATTTTCTTGAATGTGGGTAAACATCTTTAGATTGAAG	1287
Db	817	TTCTGTGTAATAA-----	806
Qy	1288	TTTTTTCTCTAGCCTCTTTTAGTCTGTCAATTTGAAGATAGATATTTCTTTACATCTGATTT	1347
Db	805	-----CCTTTGAAGGCTGGATTCGTGGAGAGATAATGTGGAAATGGTTTT	759
Qy	1348	TATCTTPAGAATGCTTTCTTTCTCCAACTATTGTGACAGAAAGTTTTTCTTAAGTCAGTA	1407
Db	758	GTGCT--GGAATACTTTGTTCTCCATCTATGTAATTCAGAGCTTTGGCTGGGTATAGTA	701
Qy	1408	GTCTGCCCTGACATCTGTAGTCTCTTGAGTCTGTAGCACATCTGTGCAGGGCCITCTTA	1467
Db	700	GCTGGGCTGGAAATTTGTGTCCTTAGTCTCTGTATAACATCTGTCCAGGATCTTCTCGG	641
Qy	1468	CATTTTGAGTTTCTATTGGAAAAGTCAGGTGAATTTCTAATACATCTGCCTCTTATATGTT	1527
Db	640	CTTTCATAGTCTCTGGTGAAAA--TCTGGTGTAAATCTGATAGGCTCGCCTTTGTATGTT	582
Qy	1528	AATTGTCCTTTTTCCCTTGACATCTTTTAAATTTCTTTTGTCTCTATCTATTTTAGTGA	1587
Db	581	ACTT-GACCTTTTCCCTTACTGCTTTAGTATTCATCTTATTTAGTGCACTTGTGT	533
Qy	1588	TTTGATATTATGCACTGGGAGATTTCTTTTCCGGTCCAACTCTATTTGGTGTTTTGTGA	1647
Db	522	TCTGATTTATGTGTCGGAGGAAATTTCTTTCTGCCCCAGTCTATTTGGAGTTCGTGA	463
Qy	1648	TGCTTCTGTACCTTGATAGGCATCTTTTCTCAAGTTAGGAAATTTTCTTTTGTGTT	1707
Db	462	GGCTCTCTGTATGTCATAGGCACTCTTTCTTTAGATTGGGAAGTTTTTCTTTTAAATAA	403
Qy	1708	TTTTCTTGAAAAATTTTTCCCTGCTTTTGACCCTG-----CCTTCTTCCCCTT	1753
Db	402	TTTGTGAGAGTGTGTGCGCCCTTTGATTTGAAAACTTCTCATCTCATCCACTCCTAT	343
Qy	1754	CCTCTATTCCTTGGTTTTTGATAGTGTCTCTGGCTCTCTGGATGTTTTATGCCCTGGAT	1813
Db	342	TATCCGTAGGTTGGTCTTCTCATGTGTGCTCGGATTTTCCCTGGATATTTTGAGTTAGGAT	283
Qy	1814	TATTTTAGACTTAAACATTTTCTTGACCAAGGTATCCATTTCTCTATCTGTGCTTTCACT	1873
Db	282	CTTTTTCATTTTCCATTTTCTTGATTTGTTGCGGATGTTCTCTATGGAATCTTCTGC	223
Qy	1874	GCTTGAGATCTCTCTCTATCTCTTGATTTCTGTGTCAGTGAGGCTGTCTCTGAGGTTCC	1933
Db	222	ACCTGAGATCTCTCTTCCATCTCTGTATCTGTGTGTGATGCTCAAAATCTACGGTTCC	163
Qy	1934	TG-TTGGGTTCTTAATTTTTTCAATTTCCAGATTTCCCTCAGTTTTGGGTTTTGTTTATAA	1992
Db	162	AGATTCTTCTTAGGGTTTCTATCTCCAGTGTGCCICACTTTTGASITTTCTTTATGT	103
Qy	1993	TTCTATTTCCATTTTCAGTCTGAGAAATGTTTATCTCATTTTCTCCAGATATTACATTT	2052
Db	102	GTCTACTTCCCTTTTAGTCTAGTATGTTTTGTTCTATTTCCATCACTGTTTGTATGT	43

Query Match	22.98;	Score 521.598;	DB 1;	Length 1644;
Best Local Similarity	62.64;	Pred. No. 2.1;		
Matches 1060; Conservative	0;	Mismatches 520;	Indels 114;	Gaps 17;
QY	399	AGATTTAAGCTGGTGGTGCAGATAGACATAGAGTATATTATTCAAATGCTCTTTATCTG	458	
Db	1634	AGCCTTAGTCATGGTGATCTGATAGGATACATGGGCAATTTTCAATATTTTGAATCTG	1575	
QY	459	TCGAGACTTCTTTGTTTGGAAATATGTAATCAATTTTGAGA--GTTTCATAGGGTCT	516	
Db	1574	TTGAGGCTGTTTGTGACCTATTATGTGCTCAATTTTGGAGAAGGTACCATCAGGGTCT	1515	
QY	517	GACAAGAAGTACAGTC--TTTGTGTTTGTGAAATAGTCTGTAATATATCTCT--AGGTCC	574	
Db	1514	GAGAAGAAGTATATCCCTTTTGTTTTAGGATAAAATGTTCTGTAATATATCTGTCAGATCC	1455	
QY	575	ACTTGTGTTATGACATCAGTTAGTCCAGCATTTCTCTGTTTCGTTTGTGTTGATGA	634	
Db	1454	ATTGTTTCATCACTTCTCTTAGTTTCAGTGTGTCCTCTGTTAG--TTTCGTTTTCATGA	1396	
QY	635	CTTAACTGTTGGAGAGAATGGGTATTGAAGTAGCCCACTATCTGTGTGAGGT--CAAT	693	
Db	1395	TCTGTCCATTGGTGAAGTGGTGTGTTGAAGTCTCCCACTATTATTGTGTGAGGTGCAAT	1336	
QY	694	ATGTGATTTTAGCTGTAGCTGTGCTCTGTTTATGAACCTGGGTGCACATCTGTTTGGTGC	753	
Db	1335	GTGTGCTTTGAGCTTTACTAAAGTTTTTTTAAAGATGGCTGCCCTGTATTTGGAGC	1276	
QY	754	ATAGACATTAAGAATTGCCATGTCCTCTGGTGGATTTT--CCTTTGAGCCATGTAGTA	812	
Db	1275	ATAGATATTTCAGAAATTGAGAGTTTCCCTCTGGAGGATTTTACCCTTGTATGATGAGAACGAAGCG	1216	
QY	813	TTCTTCCCAATCTCATCTGCTTAGTTTGGTGTTAAGTCTAT--TAGTCAGATATTAAAAAT	871	
Db	1215	CCCTCCTCTGCTTTTTTTTGATGACTTTGGGTTTGGAAGTCAATCTATCAGATATTAGGAT	1156	
QY	872	GACTGTATCGGCTTGCTCTTAGGGCCATTTGCTTAGAATATC--TTTTCATCCTTTTAC	930	
Db	1155	GGTACTCCCACTGTGTTCTTCTATACCAATTTGCTTGGAAATGTGTTTTCAGGCTTTTAT	1096	
QY	931	TCTAAGGTGATGCTATCCAT--GCTAGTGTCTTTTTTGGATGCAGCAGTAGGATGG	987	
Db	1095	TCTGAGTAGTGATGATCTTTTCTCTGAGATGTCTCTGTAAACAGCAAAATGTTGG	1036	

QY	1948	TTTTTTCATTC	CCAGATTTCC	TTCCTTCAGCTT	GGGTTTTGTTT	TATTAATTC	TATTC	CACTTT	2000
Db	1667	GATTC	TCTCT	CACTATTG	CACTCTCT	TTGGGTTTCTTT	TATGTGCT	CACGTCCTTT	1726
QY	2008	CAGGTCCT	GAAATG	TTTTTACT	CATTTTCT	CCAGATTTACA	-----	TTTTTCAT	2063
Db	1727	TAGGTC	TAGTATG	ATTTTGT	TTTATTC	ATCATCCCTG	TTGGAGT	TTTCTCTGTTTT	1786
QY	2064	CTTTAA	TGGATT	TATTCATT	CTCTCT	CAAGGACCTTT	TATGAAT	CATATAAA	2120
Db	1787	CTTTAAGG	ACTCTAC	CTGTTGGT	TGTGTTCT	CTCTGTTTTT	CTTTAAG	CACTTGA	1843

RESULT 10
AK046811/c

AK046811 4832 bp mRNA linear HTC 20-SEP-2003
 LOCUS Mus musculus 10 days neonate medulla oblongata cDNA, RIKEN
 DEFINITION full-length enriched library, clone:B330013D20 product:unknown EST,
 full insert sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE

1	Carninci, P. and Hayashizaki, Y.
AUTHORS	High-efficiency full-length cDNA cloning
TITLE	Meth. Enzymol. 303, 19-44 (1999)
JOURNAL	
MEDLINE	99279253
PMID	10349636

PUBLISHED REFERENCE

ALL AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, K., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL
MEDLINE

RECEIVED 2043337
 PUBMED 11042159
 REFERENCE 3

AUTHORS

Konno, H., Akiyama, J., Nishi, K., Kitsuami, T., Tashiro, H., Itoh, M., Sumi, N., Iehi, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system-394-format sequencing pipeline with 384 multipillar sequencer

Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL
MEDI.TNE

MEDLINE
20330913
PUBMED
11076861

REFERENCE

AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409, 685-690 (2001)

REFERENCE

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE

JOURNAL
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
REFERENCE
6 (bases 1 to 4832)

REFERENCE AUTHORS

Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawah, J., Koijima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,

Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Havashizaki, Y.

Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan. (E-mail: genome-res@scg.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

RNA-seq library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
Riken: 81-3-5083-5246/

URL: <http://genome.gsc.riken.go.jp/>
URL: <http://fantom.gsc.riken.go.jp/>.

 Location/Qualifiers |


```

Db 1290 TTT-----CCTGTTTAAACATCTGAGGTATCCATTCTTCCATCATGTCTTTAATGCCT 1238
Qy 1878 GAGATCTCTCTCTATCTCTGTTATCTGTCAGTGAGGCTTG-TCTCTGAGGTCCTGT 1336
Db 1237 GAGATCTCTCTCTCTGATCTCTGATCTATGTTGAAGGCTTGCCTCTGAGGTCCTGT 1178
Qy 1937 TGGGTTCTTAATTTTTCATTTCCAGATTTCCITTCAGTTTGGGTTTGTGTTTATTAATCT 1996
Db 1177 TCCAATCCCTAAATTTTCATTTCCAGATTTCCCTCAGCTGGGTTTCTTGTGTTACTCT 1118
Qy 1997 ATTTC-ACITTCAGGTCCTGAATGTTTACTCATTTTCTCCAGCATTTACATTTTC 2055
Db 1117 ATTACCAACTTCAGACCTTGAAGTGTGTTTACTCAATTCCTTCCACATGTTGTGTCATA 1058
Qy 2056 ATAGGTTCTTAAATGATTTATTCATTTCTCTTCAAGGACCTTTTATGAATTCATAAA 2115
Db 1057 GATG-----CAAGGATTTTATTCATTTCTCTCTAAGACTCTCATATTCATAA 1004
Qy 2116 ATGATGTTAAGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2174
Db 1003 GGTATTTTAAAGGCTTTTCTTATGTTTCAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 944
Qy 2175 AATAGGCT-----TTTACAGGACATATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2226
Db 943 TGTAGGTTACTAAGCTCTAGTGAGACATATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 884
Qy 2227 GCTTTGGCATATAGAGCTGAGTTGG 2254
Db 883 ATGGCAGCATCTAGGCATCTGGGACTGG 856

RESULT 11
AK045866
LOCUS
DEFINITION
Mus musculus adult male corpora quadrigemina cDNA, RIKEN
full-length enriched library, clone:B230315G06
product:unclassified, full insert sequence.
AK045866
VERSION
AK045866.1 GI:26091171
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
PUBMED
10349636
REFERENCE
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
PUBMED
11042159
REFERENCE
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, Y.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
Yoneda, Y., Ishikawa, T., Okawa, K., Tanaka, T., Matsuyama, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multipillar sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
PUBMED
11076861
REFERENCE
4
The RIKEN Genome Exploration Research Group Phase II Team and the

```

```

PANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3375)
ADACHI, J., AIZAWA, K., AKIMURA, T., ARAKAWA, T., BONO, H., CARNINCI, P.,
FUKUDA, S., FURUNO, M., HANAGAKI, I., HARA, A., HASHIZUME, W.,
HAYASHIDA, K., HAYATSU, N., HIRAMOTO, K., HIRAO, T., HIROZANE, T.,
HORI, F., IMOTANI, K., ISHII, Y., ITOH, M., KAGAWA, I., KASUKAWA, T.,
KATO, H., KAWAI, J., KOJIMA, Y., KONDO, S., KONNO, H., KODA, M.,
KOYA, S., KURIHARA, K., MATSUYAMA, T., MIYAZAKI, A., MURATA, M.,
NAKAMURA, M., NISHI, K., NOMURA, K., NUMAZAKI, R., OHNO, M., OHSATO, N.,
OKAZAKI, Y., SAITO, R., SAITO, H., SAKAI, C., SAKAI, K., SAKAZUME, N.,
SANO, H., SASAKI, D., SHIBATA, K., SHINGAWA, A., SHIRAKI, I.,
SUGABE, Y., TAGAMI, M., TAGAWA, A., TAKAHASHI, F., TAKAKU-AKASHI, S.,
TAKEDA, Y., TANAKA, T., TOMARU, A., TOYA, T., YASUNISHI, A.,
MURAMATSU, M. and HAYASHIZAKI, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
FEATURES
Source
Location/Qualifiers
1..3375
/organism="Mus musculus"
/mol_type="mrna"
/strain="C57BL/6J"
/db_xref="FANTOM DB:B230315G06"
/db_xref="MGI:2411246"
/db_xref="taxon:10090"
/clone="B230315G06"
/sex="male"
/tissue_type="corpora quadrigemina"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
misc_feature
1..3375
/notes="unclassified"
Query Match 22.1%; Score 503.498; DB 1; Length 3375;
Best Local Similarity 62.4%; Pred. No. 1.3;
Matches 1091; Conservative 0; Mismatches 490; Indels 167; Gaps 21;
Qy 384 TTGTTGTTGTTATCTAGATTAAAGTGTGCTGTCAGATAGGACATAGATATTATTTCA 443
Db 1 TAGTTGTTGAAGCCAGCTTTAATCCATGCTGATCTGATAAAACACCCAGGGGTCATTTC 60
Qy 444 ATTGTCCTTTTATCTGTCGAGACTGCTGTTGTTTGAATATGTTTCAATTTTGGAGAGT 503
Db 61 ATCTTTCTGTAUCTTTTGGGCTTACTCTGTCCTGATCATATGTTCAATTTAGAGAAG 120
Qy 504 TTCATAGG--GTGCTGACAGAGGTACAGTC--TTTGTTGTTTGGTGAATAGTCTGTAA 560
Db 121 ATTCTGGGATGTAAGGAGGAGGATATTTCTTTGTTTGTGAGTGAATGTTCTGCAA 180
Qy 561 ATATCTC-TAGGTCACCTTGGTTTATGACATGAGTTAGTCCAGCATTTCTCTGTTTCT 619
Db 181 ATATCTCTAAGTCCTTTGGTTTCATACAT--GTTAGTTTCATTTCTCTATTATG- 237

```

QY 620 TTTTGTGAGATGACCTAACTGTT--GGAGAGAAATGGGGTATTGGAAGTAGCCCACTATC 677
 Db 238 CTTCTGTTTGTGATGACCTGTCATTTGAGGTGAGAGTGGGGTGTGAAGTCTCCCAATATT 297
 QY 678 TGTGTGTGAGG--TCAATATGTGATTTTAGCTGTAGCTGTGCTGTGTTTATGAACCTGGGT 736
 Db 298 ATTGTGTGAGGTCAATGTGTGATTTGTGCTTTAGTAAAGTTTCTTTTACAAGTGTGAGT 357
 QY 737 GACATGTGTTTGGTGCATAGACATTAAGAAATGAAGTCCCTCTTGGTGAATTTCCCT 796
 Db 358 GCCCTTGCATTTGAGTACAGATGTTCCAGAAATGAGATATCATCTTGGTAGAATTTTCCCT 417
 QY 797 TGATGCCATGTAGTATTTCCCAATCTCATCTGCTTAGTTTTGGTGTAAAGTCTA--TT 855
 Db 418 TGATGCATATGAAGTGTCTTCTCACTCTTTTGAATACCTTTGGTGTAAATCCATTT 477
 QY 856 AGTCAGATATTAAGATGACTGTATGGCTTCTTTAGGGCCATTTGCTTGAATATC- 914
 Db 478 TATTAAATATTAGATTCACTACTCCAGCTTGTCTTGGGTGGTGGCTGGAAACCT 537
 QY 915 TTTTCCATCCTTTTACTCTAAGGTGATGCTA---TCCATGGTAGGTGTCTTTTGGGA 971
 Db 538 TTTTCCAGCCCTCTATGCTGAGGTAGTGTCAACTTTGTGTCTGAGATGTTTCTTGTGA 597
 QY 972 TGCAGCAGTAGGATGGATCTTGTGTTTCATATCCATCTGTTTACCAGATCTTTTTTCTAG 1031
 Db 598 TGCAGCAGATGATAGATTCTGTTTGTGATCCAGTCTTTTAGCTGTGCTTTTATCA 657
 QY 1032 AGAAATTAAGATCAATGAGTCATGATGTTGAGAAATTAATCAATGAGCAGTGTGTTGCGAT 1091
 Db 658 GAGAACTGAGTCCCACTGA---TGTTGAGATGAGAGATATTAAATGATCAGTGTGTTGTATT 714
 QY 1092 TCTGTGTTATCTGCACTGTGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1151
 Db 715 TCTGTGTTATTTGATGTTG----- 733
 QY 1152 TGT 1211
 Db 734 -----GGTATGT 755
 QY 1212 CCCCTTTTGTGTTTGGCCCTGGATATTATTATTATTCATATTTCTGTAATGTGGGTAA 1271
 Db 756 -----TCTTTTCTGTGAAATGAGTAATTTCTTGTGTCTCTTGGGGGTAGTTAC 804
 QY 1272 CATCTTTAGATTGAAGTTTTTCTCCTAGCCCTTTTAGGTCTGATTTGAAGATAGATAT 1331
 Db 805 CTTCTCTGTGTTGGAT-----GTAGATAT 828
 QY 1332 TCTTTACATCTGATTTTATCTAGAAATGTCTTTCTTCTCACAATAT-TGTGACAGAAAG 1390
 Db 829 TGTTTAACTTTGTTTATATCAATGAGATATTTTGGTTTCTCATCTATCAGTGAATTAAGAG 888
 QY 1391 TTTTCTTAAGTGCAGTACTGTGGCTGACATCTGTAGTC--TCTTGAGTCTGTAGACACA 1448
 Db 889 TCTTGTGTTGTTATAGTAGTCTGGCTGCCATCTGTGTTCTTTTAYAGGCTGTAGACACA 948
 QY 1449 TCTGTGAGGCGCTCTTTPACATTTTGTAGTTTCTAATGAAAGTCAAGGTGAATCTTAAT 1508
 Db 949 TCTGTTCAGGATCTTCTGGCTTTTATATAGTCTCTGTGAGAAATCAG---GTAGTCTGTAT 1005
 QY 1509 ACATCTGCTTTATATGTTAATGCTTCTTTTCTCCCTTGATCTTTTAAATATCTTTCTT 1568
 Db 1006 AAGTCTGCTTTTATAGGTTTCTTGGCCCTTTTCTCTGTGAGCTTTTAAATATCTTTCTT 1065
 QY 1569 TGTCTTACTTTTGTAGTGAATTAATATGCACTGTGGGAGTGTCTTTTCCGGTCCA 1628
 Db 1066 TGTCTGTACATTTAGTGTGTTTGTATATGTCGAGGAAGATTTCTTTTGTGTCATA 1125
 QY 1629 ATCTATTGTTGTTTGTATGCTTCTGTACCTTGTACCTGTAGCATCTCTTCTCAGGTAG 1688
 Db 1126 ATCTATTGTTGTTCTGTATGCTTCTATGAGTGTGTTTGTAGTATCTCTTT-TTTAGGTG 1184

QY 1689 GAAATTTTCTTTTGGTTTTCTTGAAATATATTTTCCCTGCTTTTGGACCTG----- 1740
 Db 1185 GGAAGTTTTTCTTCTATGATTTTGTGAAGATATGTTCTGTCCCTTTGAACTGGGATCTT 1244
 QY 1741 -----CCTTCTTCCCTTCCCTCTATTTCTTGGTTTTTGGCATAGTGTCTCTGGCTTCCCTG 1795
 Db 1245 CACCTTCTTCTATCTATTTATTTCTTAGATTTGCTCTTTTCTATAGTGTCCCAATTTTCCCTG 1304
 QY 1796 GATGTTTATGCTGAGTATTTTATGACATTAACATTTCTTTGACCAGGTATCCATTTCC 1855
 Db 1305 GATGTTCTGAGTTAGAACACTTTTATGATTTGGCATTTTCTTTGACATCTATCAATTTA 1364
 QY 1856 TTTATCTTGTCTTCACTGCTGAGATTTCTCTTCTCTATCTCTTGTATTTCTGTCTGAGTGA 1915
 Db 1365 TACTATGGGATCTTCTACACCTGAGATTTCTCTCTTC-ATCTCTTCTATTTTGTGTTAAT 1423
 QY 1916 GTTGTCTCTGAGGTCTCTGTTGGGT--CTTAATTTTTCATTTTCCAGATTTCTCTCAG 1973
 Db 1424 GATTCATTTACTAGTCTGTTCTCTTCTCTAGATTTTCCAACTCCAGGTTGCTCCA 1483
 QY 1974 TTTGGTTTTGTTTATTAATCTTATTTCCACTTTCCAGTCTCGAAATGTTTACTCATTT 2033
 Db 1484 TTCAGTTTTTCTCAITGCTTCTATTTCTGTTTTCAGTCTTGAACAATTTTATCATTT 1543
 QY 2034 TCTCCCACTATTTACATTTTCCATAGTCTTCTTAAATGGAATTTATCAATTTCTCTCAA 2093
 Db 1544 CTTTCCACTGTTTGTATGTTCTTTTCTATTTCTTTAAGGATTTCTTCTCTCTTAA 1603
 QY 2094 GGACCTTT 2101
 Db 1604 GGACTTCT 1611

RESULT 12
 AK034168
 LOCUS
 DEFINITION Mus musculus adult male diencephalon cDNA, RIKEN full-length enriched library, clone:9330161A6 product:unclassified, full insert sequence.
 ACCESSION AK034168
 VERSION AK034168.1 GI:26083804
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE 2
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159
 REFERENCE 3
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsu, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861

Query Match		20.1%; Score 457.898; DB 1; Length 2108;
Best Local Similarity		62.0%; Pred. No. 3;
Matches 1034; Conservative		0; Mismatches 496; Indels 137; Gaps 21;
QY	482	TATGATATCAATTTTGAGAG- GTTTCATAGGGTGTGACAGAGGTACAGTCT-TTGT 538
DB	1973	TATATGTCATTTTGAAGAAGGTTCCATGAGGTGTTGAGAAGAAGATATATTTCTATTGT 1914
QY	539	GTTTTGTGAAATAGTCTGTAAATATCT-CTAGTCCACCTTGGTTTATGACATCAGTTAG 597
DB	1913	TTTAGTATAAAGTCTGAGAAATATCTGTATAATCCATTGGTTTCAACATCTCTATTAC 1854
QY	598	CTCCAGCATTTCTCTGTTTCTGTTTTTTTGTGAGATGACCTAACTGTGTGGAGAAATGGGG 657
DB	1853	TTTCCACTGTGCTCTGTATAGTTTC-TGTTTCCAGGATCAGTCTCTGTGACAGAGTGGGA 1795
QY	658	TATTGAAGTACCCACATCTGTGTGTGT- GAGGTCAATATGATTTTATGCTGTAGCTGTG 716
DB	1794	TGTTGAAGTCTCCACATATATTTGTGTATAGTCAATGTGTGCTTTGAGCTTTACTAATA 1735
QY	717	CTTGTGTTATCAACTTGGGTGACATTTGT- TTTGGTGCATAGACATTAAGA 766
DB	1734	TTTCTTTAATGAATGTGATGCCCTTGCATTAATAAATATCTTGGAGCATAGATATTCAGA 1675
QY	767	ATTGCAATGCTCTTGTGTGATTTTCTTTGTAGTCCCTATGTAGTATCTTCCAAATCTC 826
DB	1674	ATTGAGAGTTACCTTAGTAGATTTTACCTTTGATGAGTATGAAGTATCCTTCCCTTATCTT 1615
QY	827	ATCTGCTTAGTTTGGGTTTAAAGTCTA-TTAGTCCAGATATTAATAATGACGTATCGGCTT 895
DB	1614	TTTTGATACTTTAGTTGAAGTCGATTTTATTCATATTAAGATGAGTACTCCAGCTT 1555
QY	886	GCCTCTTAGGCCCAATTTGCTTAGAATATC-TTTTCCATCCCTTTTACTTAAGTGTATGTC 944
DB	1554	GTTTCTTGGGACCAATTTGCTTGGAAATTTGTTTTCCAGCCCTTTTACTCTGAGGTAGTGTG 1495
QY	945	TATCCATGGTA -GGTGTCTTTTGTGATCCACAGTAGAGTCACTTGTGTTTTCATAT 1002
DB	1494	TGCTTTGCTACTGAGTGGTCTGTATGACAGAAATGCTGGGCTCTTTTATGTATC 1435
QY	1003	CCATCTGTTACCCAGTATCTTTTCTAGAGAAATTAAGATCATTTAGTCAATGATGTG 1062
DB	1434	CCAGTCTTTTAAATATATCTTTT- TATTAGGAAATGAGTCCATTCATATTA 1382
QY	1063	AGAAATATCAAGACGAGTGTGTGGATCTTGTGTATCTTGCACTTGTGAGTGTGTG 1122
DB	1381	AGAGATATTAAGTAAAGTAATATTTGTTCTCTGTATTTTGTGTAGAGTTGGAAT 1322
QY	1123	GT 1182
DB	1321	CTGTTCACTGGCTATCTCTTTAGGTTG----- 1291
QY	1183	GTCTGTGTCTTGTGTGTGTGTGTCTCTCCCTCTTTTGATTTTGGCCTGGAATATTT 1242
DB	1290	-----TTGAAGATTA 1280
QY	1243	ATTATTCATATTTCTTGAATGTGGTAACATCTTAGATTGAAGTTTCTCTCCAGCT 1302
DB	1279	TTTCTGTCTTTTCTAGGCAATAGTTTCCCTCTCTGTTGTTGAGTTTCTCATTTATGT 1220
QY	1303	TCTTTT---AGTCTGCAATTTCAAGATAGATATTTTCTTACATCTCATTTTATCTAGAA 1358
DB	1219	CTTTTGAAGGCTGGACATATGGAAGATATATTTGTTTAAATTTGTTTCTCATGGAAT 1160
QY	1359	GTCTTTCTTCTCCAACTATTTGTGACAGAAAGTTTTTCTTAAGTGCAGTGTCTGGCTGA 1418
DB	1159	ACCTTAGTTTCTCCATCTATGATTAATTGAGAGTTTTTGTGGGTATAGTAGCTAGGCTGG 1100
QY	1419	CATCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGACAGGCTCTTACATTTTGTAGTT 1478
DB	1099	CATTGTGTCTCTTAGGCTGTATGACATCTGCCAGGATCTTCT--GGTTTTCATAA 1042
QY	1479	TCTATTGGAAGAGTCAGGTGTAATTTCTAATAACATCTGCTTTTATATGTTAATGTGCTTT 1538
DB	1041	TCCTCTGGCAAGAAGTTTGGTGTAAATTTGATAGGTCTGCTTTATATGTACTTCACCTTT 982
QY	1539	TTTCCCTTGCATCTTTTAAATATCTTCTTGTCTTCTATACATTTTAGTATGATTATTA 1598
DB	981	TTT-CTTACTCTCTTTTAAATATCTATCTTGTGTCATTTGTCATTTGTTTGTATATTA 923
QY	1599	TGCACCTGTGGGAGATTTCTTTTCCGGTCCAATCTAATTTGGTGTGTTTGTATGCTTCTGTA 1658
DB	922	TTGACAG-GAGGAATTTCTTTTCTGCTCAATCTATTTGGAGTCTCTATAGGTTTCTTGT 864
QY	1659	CTTGTATAGGCACTCTTCTCAAGGTTAGGAATTTTCTTTTGGTGTGTTTCTTGTGAAA 1718
DB	863	TGTTCAATGGCACTCTTTTTCAGGTTAGGAAGTTTCTTCTCTAAATTTTGTGGAAGA 804
QY	1719	TATTTTCCCTGCTTTTGACCT-SCCTTCTTCCCTCTCTCTATTC-----T 1764
DB	803	TATTTACTGGCCCTTTAAGTTGGAATCTTCACTGTCTCTATACCTACTATACAGGT 744
QY	1765	TTGGTTTTGCAATGCTCTGGCTCTCTGAGTGTGTTTATGCCGTTAGTATTTTAG-AC 1823
DB	743	TTGCTCTTCTCATTTGTCTCGAGTCTCTGATGCTTTGAGTTAGGATCATTTTTCAT 684
QY	1824	TTAAACATTTTCTTGACCAAGGTATCCATTTCTTCTATCTTCTCTCTCACTGCTCAGATT 1883
DB	683	TTTGCATTTCTTGTATGCTGTGCAATGTTTCTGTTGTAATCTCTGACTCTGAGATT 624
QY	1884	CTCTCTCTATCTCTTGTATCTGTGAGGAGTGTCTCTGAGGTTCTGTG- TTGGGTT 1942
DB	623	CTCTCTCTACCTCTGTATCTGTGTTGTAATGCTGATCTATGCTCTGCTCTGACTCTTT 564
QY	1943	CTTAATTTTCTCATTTCCAG---ATTTCCTTCAGTTGGGTTTGTGTTTATTAATTTCTAT 1998
DB	563	CTTAGATTTCTATCTCCAGATTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 504
QY	1999	TTCCACTTTTCAGTCTCTGAAATGTTTACTCATTTTCTCTCTCTCTCTCTCTCTCTCTCT 2054
DB	503	TTCCATTTTGTAGTGTGTAATGTTTATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 444
QY	2055	CATAGTGTCTTTTAAAGATTTTATTCATTTCTCTCTCTCAAGGACCTTT 2101
DB	443	CTGTAATTTCTTAAAGGATTTTATGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 397
RESULT 14		
AK048889		3179 bp mRNA linear HTC 20-SEP-2003
LOCUS		Mus musculus 0 day neonate cerebellum cDNA, RIKEN full-length
DEFINITION		enriched library, clone:C230078P05 product:carbonic anhydrase 5a, mitochondrial, full insert sequence.
ACCESSION		AK048889
VERSION		AK048889.1 GI:26093182
KEYWORDS		HTC; CAP trapper.
SOURCE		Mus musculus (house mouse)
ORGANISM		Mus musculus
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS		Carninci, P. and Hayashizaki, Y.
TITLE		High-efficiency full-length cDNA cloning
JOURNAL		Meth. Enzymol. 303, 19-44 (1999)
MEDLINE		99279253
PUBMED		10349636
REFERENCE		
AUTHORS		Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE		Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL		Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE		20493374
PUBMED		11042159
REFERENCE		
AUTHORS		Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,

Db 2390 CCTCTGTAGCGCTGGATAAGTGAAGAAATATGTTTAAATTTCTTTTTCGTTGGGAATA 2449
Qy 1360 TCTTTCTTTTCCAACTATTGTGACAGAAAAGTTTCTTAAGTCAGTAGTCGCGCTGAC 1419
Db 2450 TACTGTTTCTCCATCATGTGATGAGAAATTTGCTGGGTATAGTAGCTGGCAGAC 2509
Qy 1420 ATCTGTAGTCTCTGGAGTCTGTAGCACATCTGTGAGGCGCTTCTTACATTTTGAGTTT 1479
Db 2510 ATTGTGTTTCCGATAAGTCTGTATGACATTTTCCAGATCTCTCGACTTTTATTGTGT 2569
Qy 1480 CTATTGAAAAGTCAGGTGAATCTTAATACATCTGCCTTTATATGTTAATGCTGCTTT 1539
Db 2570 CTGTT-GAGAAGTCGTGTTAAATTTGATAGGTTGACTTTATATGTTACTT-GACCTTT 2627
Qy 1540 TTCCCTTGCAATCTTTAAATATCTTTCTTTGTTCTATACATTTTAGTGATTTGATTAAT 1599
Db 2628 TTCTCTTACTGTTTAAATATCTGTTCTTTGTTCTGTGATTTAAATGTTTGAATTTT 2687
Qy 1600 GCACTGTGGGAGTTCTTTTCGGTCCAACTATTTGGTGTGTTTGTATGCTCTTGTAC 1659
Db 2688 GTGACATGAATAATTTATTTCTGGCAAACTATTTGGTGTGTTCTGTAGGCTCTTGTAT 2747
Qy 1660 CTTGATAGGATCTCTTTCTCAAGTTAGGAAATTTTCTTTTGGTTTCTTGAAT 1719
Db 2748 GTTCATGGCATCTATTGTTTGTAGTTAGGAAATTTTC--TATGAATCGGTTGAAGAT 2804
Qy 1720 ATTTTCCCTGTTTGAACCTCTTCCCTTCTCTCT-----ATTCCCT 1765
Db 2805 GTTTTCTGACCTTTGAACCTGGAATCTCCACTCTCTTTATTTCATTAATTTCTTAGCT 2864
Qy 1766 TGGTTTTGTCATAGTCTCTGCTTCCCTGGATGTTTATGCTCGGATTTATTAGACTT 1825
Db 2865 TGATCCTTTTCATCGTGTACTGAATTTCTGATGTTTGTAG--TTAACTTTTGCACCT 2921
Qy 1826 AACATTTCTTTGACCAAGTATCCATTTCTTCTATCTTCTCTCTCTGCTGAGATCT 1885
Db 2922 TGTATTTTCTTTGATTTGTTGTCATATCTTCTATGTTATCTTCTGCTGAGATCT 2981
Qy 1886 CTCTTCTATCTTGTATTTCTGTCAGTGAGCTTCTCTGAGGTTCTCTGCTGCTGTTCT 1945
Db 2982 GTTGTTTATCCCTCATATTCGTTGGTGTGATGCTGATCTGATTTCTGATCTTCTCC 3041
Qy 1946 AATTTTTCATTTCCAGATTTCTCTAGTTTGGTTTGTGTTTATTAATTTATTTCCACT 2005
Db 3042 TGGTTTTCCATCTCCAGGTTGCTTCTATTTGTAATTTCTTTATTGTTCTTCTTCCATT 3101
Qy 2006 TTAGGTCCTGAATGTTTACT 2028
Db 3102 TTTAGATCCTGAATATGTTGCT 3124

RESULT 15
AK079106
LOCUS Mus musculus adult male diencephalon cDNA, RIKEN full-length
DEFINITION enriched library, clone:9330197K07 product:unknown EST, full insert
sequence.
ACCESSION AK079106
VERSION AK079106.1 GI:26098277
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
AUTHORS High-efficiency full-length cDNA cloning
TITLE Meth. Enzymol. 303, 19-44 (1999)
JOURNAL 99279253
MEDLINE 10349636
PUBMED
REFERENCE 2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsushita, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 6
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanakawa, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koyama, S., Kurahara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, P., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration and Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
FEATURES
Location/Qualifiers
1. 4502
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="FANTOM:DB:9330197K07"
/db_xref="MGI:2398672"
/db_xref="taxon:10090"
/clone="9330197K07"
/sex="male"
/tissue_type="dienecephalon"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
misc_feature
1. 4502
/note="unknown EST (GB|BB455937, evidence: BLASTN, 99%,


```
match=663)"
```

Query Match 18.9%; Score 430.098; DB 1; Length 4502;
Best Local Similarity 58.4%; Pred. NO. 1.9;
Matches 1038; Conservative 0; Mismatches 579; Indels 151; Gaps 18;

Query Match	18.9%; Score 430.098; DB 1; Length 4502;	
Best Local Similarity	58.4%; Pred. No. 1.9;	
Matches 1038; Conservative 0; Mismatches 579; Indels 161; Gaps 18;		
Qy	532	TCCTTGGTTTGGTGAATAGTCGTAAATATCT-CTAGGTCCACTTGGTTTATGACAT 590
Db		
S	5	TTTTTTTTTTAGGATAAAATGTTCTATAGATATCTGTTAAATCCATTTTGCCATAACTT 64
Qy	591	CAGTTAGCTCCACATTTCTCTGTTTCGTTTTTTTGTGAGATGACCTAACTGTTGGAGAG 650
Db		
65	ATGTAGTTTCACTGNGTCTTTGTTTTAG-TTTCTGTTTTCCATGATCTGTTCAATGCTGAG 123	
Qy	651	AATGGGGTATTGAAGTAGGCCA---CTATCTGTGTGGAGGT-CAATATGTGATTTTAGC 706
Db		
124	GGTGGG-TGTTGAAGTCTCCAAATGTTATTTACTGTGTGAGGTACAAATGTGTCTTTGAGC 182	
Qy	707	TGTAGCTGTCTGTTTTATGAACCTGGGTGACATTTGTTTGGTGCATAGACATTAAGA 766
Db		
183	TTCACTAAAGTTTCCTTTTATGAATGTGTGTGGCCTTGCAATTTGGAGCACAGATGTTGAGA 242	
Qy	767	ATTGCAATGCTCTCTTGGTGGATTT-TCCCTTGTAGTCCCTATGTAGTATTTCTTCCCAATCT 825
Db		
243	ATTGAGAGTTCTCCCTTGGTAGATTTTCCCAATGATGAGTATGAAGTGTCTCTTCTACCT 302	
Qy	826	CATCTGCTTAGTTTGGGTTTAAAGTCTATAGTTCAGATATTAATAATGACTGTACGGCTT 885
Db		
303	TTTTTGATACTTTTGGTTGAAAGTTCCTTTTATTCGATATTAGAATAGCTCTCCAGCTT 362	
Qy	886	GCTTCTTAGGGCCATTTGCTTAGAATA-TCTTTTCCATCTCTTTACTCTAAGGTGATGTC 944
Db		
363	GTTTCTTCAGACCGTTTGTCTGGAAATTTGTTGTCCATGTTTTATTCTGAGGTAATGCC 422	
Qy	945	TATCCATGTTAG---GTTGTCTTTTTTGGATCCAGCAGTAGGATGGATCTTGTTTTCATA 1001
Db		
423	TGTTTTTGTCACTGAGGTGCATTTCTGTATGAAGCAAAATGTTGGGCTCTGTTTATGTA 482	
Qy	1002	TCCATTCGTGTACCCAGTATCTTTTTCTAGAGAAATTAAGATCAATTGAGTCAATTGATGTT 1061
Db		
483	TCCAGTCTGTATTCTATTCTTTTTTATGGAGATTAA-----ATCCACTGATATT 534	
Qy	1062	GAGAAATTACAATGACGAGTGTTTGGGATTCCTGTTATCTTGTGCACTTGTGGAATGTGTG 1121
Db		
535	AAGAAATATTAAGAAATAGTATTGTGTCTCTGTTATTTGTGATGTTATTTTTATATT 594	
Qy	1122	TG 1181
Db		
595	TCTGTGGTATCTCTCTTTTGGGTTTGTTHAAGAGAG----- 630	
Qy	1182	TGTCGTGTGTGTGTGTGTGTGTCTCTCCCTCTTTTGTGATTTTGGCCCTGGAATATT 1241
Db		
631	-----ATT 633	

RESULT 16					
BC021486/c					
LOCUS	BC021486	2278 bp	mRNA	linear	HTC 04-MAR-2003
DEFINITION	Mus musculus, clone IMAGE:5344624,				
ACCESSION	BC021486				
VERSION	BC021486.1	GI:20987654			
KEYWORDS	HTC.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 2278)				
AUTHORS	Strausberg R.				
TITLE	Direct Submission				
JOURNAL	Submitted (14-JAN-2002) National Institutes of Health, Mammalian				
	Gene Collection (NGC), Cancer Genomics Office, National Cancer				
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,				
	USA				
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov				

COMMENT

Contact: MGC help desk
Email: cgapbs-@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@hghri.nih.gov
Akhter, N., Aysle, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Madero, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 55 Row: c Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis
This clone has the following problem: retained intron.

FEATURES

source

Location/Qualifiers
1. .2278
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGES:5344624"
/tissue_type="Mammary tumor, C3(4)-Tag model. Infiltrating
ductal carcinoma. 5 month old virgin mouse."
/clone_lib="NCI CGAP_Mam6"
/lab_host="DH10B"
/note="vector: pcwv-SPORT6"

Query Match 18.8%; Score 428.398; DB 1; Length 2278;
Best Local Similarity 59.0%; Pred. No. 3.7;
Matches 954; Conservative 0; Mismatches 546; Indels 116; Gaps 14;

Qy 285 AACACATCTATTTCTTGATTTCTATCTGGCTGATTTTAACTCAGTAGTGTGTTT 344
Db 1927 AAAGCTTTAAATTTCTTTATTTCTCTCTGACCAAGTTATCATCTAGTA-----TTGTTT 1873
Qy 345 GGTTTCCATAAGTTGTAAAGTTTCTGTTGTTTCTGTTGTTGTTGTTGTTATCTAGATT 404
Db 1872 AGCTTCCACATGATGCGCTTGATTTATTTATGTTGTTATGATATC---AGCCTT 1816
Qy 405 AAGCTGTGGTGGTCAGATAGACATAGATATTTTCAATTTGCTTTTATCTGTGAGA 464
Db 1815 AGTCCATGGGATCTGATAGATGATGATGGATTTTCAACATTTCTATATCAGTTGAAG 1756
Qy 465 CTTCCTTTGTTTGAATATGATATTCATTTTGGAGAGTTTCATAGGTCCTGACAGAA 524
Db 1755 CTTGCT-----GATTATATGTCATTTTGGAGATGT-ACCATGAGGTGCTAAGAA 1705
Qy 525 GGTCAGCTTTTGTGTTTGTGTAATAGTCTGTAATAATCTCTAGGTCCACTTGTGTTA 584
Db 1704 GGTATATCTTTTCTTTTAGGATAGATGTTCTATAGATATGTTACATCCATTTGTTTCA 1645
Qy 585 TGACATCAGTTAGTCCAGCATTTCTCTGTTGCTTTTGTGTCAGATGACCTAATCTGTT 644
Db 1644 TAACTTCTGTAGTGTCAATGTATGTTG-TTGAATTTGTTTTCAGGATCTGTCCACT 1586
Qy 645 GGAGAGATGGGTATTTGAAGTAGCCCACTATCTGTGTGTG-AGGTCAATATGTGATTTT 703
Db 1585 GATGAGATGGGAGTTGAAGGCTCCCACTATATTTCTGAGCAGTGCATTTGTGTTTG 1526
Qy 704 AGCTGTAGCTGTGCTTGTGTTTATGAACTTGGGTGACATTTGTTTGTGTGATGACATTA 763

Db 1525 AGCTTTACTAAAGTTTCTTTAATGAATGTGGATGCCCTTGCAATTTGGAGCATATATGTTTC 1466
Qy 764 AGAATGCAATGTCCTCTTGTTGGATTTT-CTTTGATGCCCTATGATGATTTCTTCCAA 822
Db 1465 AGAAGTGAAGTTCACCTTGGTAGATTTTACCTTTGATGAGTATGAAATGCTTCCCTGT 1406
Qy 823 TCTCATCTCTTGTAGTTTGGGTTTAAAGTCTATTAGTCAGATATTAAATGACTGTATGG 882
Db 1405 CTGTTTCTTTGGTAACTTAGTTTAAAGTCAATTTTATCTATATCTATCTATCTCCAG 1346
Qy 883 CTGCTTCTTAGGGCATTGCTTAG-ATAATCTTTTCCATCTCTTTTACTCTAAGGTGAT 941
Db 1345 CTGTTTCTTAGGGCATTGCTTAGGAAATTTGTTCCCAACCTTTTACTCTAAGGTAGC 1286
Qy 942 GTCTATCCATG--GTAGTGTCTCTTTTGGATGCAGCAGTAGGATGATCTTGTGTTTC 998
Db 1285 ATTGTCTTTGACCTGAGTAGGTTTCTGTATGACGCAAAATGTTGAGTCTCTGTTAC 1226
Qy 999 ATATCAATCTCTTACCAGTATCTTTTCTAGAGAAATTAAGATCATTTGATGATTTAT 1058
Db 1225 AAAGCAGTCTCTTAGTCTATGCTCTTTTATTGGTGAATTTGA-----ATCCATTGAT 1174
Qy 1059 GTTGAGATTAATCAATGACGAGTGTGTTGGATTTCTGTTATCTTTGACACTTGTGAAGTGT 1118
Db 1173 ATTGAATGATATAAGGAAAGTAATTTGTTGCTTCTGTTATTTT----- 1129
Qy 1119 GT 1178
Db 1128 ----- 1129
Qy 1179 CTGTGTGTGTGTGTGTGTGTGTGTCTCTCCCTCTTTTGATTTTGGCCTGGAATT 1238
Db 1128 -TTGTTGTAGATTTGAGATTTCTGTTTGGCTATCTTCTTTAGGTGTTGTTGAAAAA 1070
Qy 1239 ATTATATTATCATATTTTCTGAATGTGGTAAACATCTTTAGATTTGAAATTTTCTCTCTTA 1298
Db 1069 TTATTTTCTTGCTTTTCTAGGTTGTGTTTCCCTCTGCTGTGTGTGTGTGTGTGTGTGT 1010
Qy 1299 GCTTCTTTTGTGCTGATTTGAAGATAGATATTCTTTACATCTGATTTTCTTTAGATTT 1358
Db 1009 ATCCCTTTGATGGCTGGAATTCATGTAAGATATCTGTGTAATTTGTTTGTCTGATGAAT 950
Qy 1359 GCTTCTTTTCTCCAACTATTCTGCAGAGAAAGTTTCTTAAGTCAGTAGTCTGCCCTGA 1418
Db 949 ATTTTGGTTTCTCCATCTAAGTAAATTTGAGAGTTTGTGCTGATATGATGATGCTGGCTGG 890
Qy 1419 CATCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGACAGGCTTCTTACATTTTGTAGTT 1478
Db 889 CATTTGATTTCTTTAGGCTATATATGACATCTGTCCAGGATCTTCTGCTTTTATAGTC 830
Qy 1479 TCTATGGAAGTCAAGTGAATTTCTAATACATCTGCTTTTATATGTTAATGTCCTTT 1538
Db 829 TCCGGT-GAGAAAGTCTGTTAATTAATAGGCTGCTGCTTCAATATCTTACTT-GACCTT 772
Qy 1539 TTTCCCTTGCATCTTTTAAATTTCTTTTCTTCTATATCTTTTGTAGTGAATTTGATTATTA 1598
Db 771 TTTCCCTTACTCTTTTAAATTTCTTTCTTTGTTGCTGCAATTTTGTATTTGATTATTA 712
Qy 1599 TGACATGTGGGAGTTTCTTTTCCGGTCCAAATCTATTTTGGTGTGTTTGTATGTTCTTGTGA 1658
Db 711 TGTGATGGGAGAAATTTTCTTCCAGTGCAAACTATTGGAGTTCTGTAGGCTTCTTTATA 652
Qy 1659 CTTGATAGGCACTCTTTTCTCAAGTTAGGAAATTTTCTTTTGTGTTTCTTTGTTGAAAA 1718
Db 651 TGTTCATGAGAAATCTTTTCTTTAGGTAGGAAATTTTCTTCAATTTTGTGTTGTTGTTT 592
Qy 1719 TATTTTCCCTGTTTGTAGCTG-----CTTTCTTTCCCTCTCTTATTTCTTTGTTT 1772
Db 591 TATTTTACTGCTTCTTTAAGGTGAAATCTTCAATTTTACCTTATTTGCTTAGGTTTCTTCT 532
Qy 1773 TGCAATAGTGTCTGCTTCTGCTGATGTTTATGCTGCTGGAATTTTGTAGACTTACATTT 1832
Db 531 TTTCAATGTGTCTGATTTCTCTGAATGTTTGTAGTAGGATCTTTTGTGATTTTGTGATTT 472

Db	2460	GTATTGTTCTTAGTTCATTGCTTGGGAAACATTTTCCAGTATTTTACTCTAAGATAA	2519
Qy	941	TGTCCTATC---CATGGTAGTGTCTTTTGTGATGCAGCAGTAGAGATGATCTTGTTTT	997
Db	2520	TGTCATCTTTGTTGCTGAGATATGTTCTGTATTACAGCAGATGATGGTTCTGTTTA	2579
Qy	998	CATATCCATTCTGTTACCCAGTATCTTTTCTAGAGAAATTAAGATCAATTGATGCTATTGA	1057
Db	2580	AGCATCCATTCTGTGTAGCTGTGGCTTTTATTGGGAATTGAGTC-----CATTTGA	2631
Qy	1058	TGTTGAGAAATTAACAATGACAGTGTGTGGGATCTTTGTATTCTTGCACCTGTGAAGTG	1117
Db	2632	TGTTGAGAGATAT---TGATCAATGATGCTACTTCTCTGTAATTTTGATGTGG-GTAGTG	2687
Qy	1118	TGT	1177
Db	2688	GTTGT	2747
Qy	1178	TCTGT	1237
Db	2748	TGTT-----	2751
Qy	1238	TATTTATATTTCATATTTTCTGAATGGGTAAACATCTTTTAGATTGAAATTTTCTCCT	1297
Db	2752	-----TTTTCTTGAGTGTAGTTAGCCCTACCTGGGTTGGAGTTTTCTCTCT	2796
Qy	1298	AGCCT--TCATTAGGCTGCAATTTGAAGATAGATATTTCTTTACATCTGTATTTTATCTTAG	1355
Db	2797	AGTATCCTCTGTAGGGCTGGGTAGCGTAATGATGTGTTTAAATTTGGTTTTGTCAATGG	2856
Qy	1356	AATGTCCTTTCTTTCTCCAACTATTGTGACAGAAAGTTTCTTAAGTGCAGTAGTCTGGCC	1415
Db	2857	AATATCTTGGTTTATCAATCTATAGTAGATTGAAAGTTTTCCTGGTATAGTAGTCTGAGT	2916
Qy	1416	TGACATCTGTAGTCTC--TTGAGTCTGTAGACACATCTGTGCAGGCGCTTCTTACATTTTG	1474
Db	2917	TGACATTTTGGTCACTTTAGGGTCTCGGACACATCTGTCCAGGCCCTTCTAGCTTTTAG	2976
Qy	1475	AGTTTCTATTGGAAAGTCAGGTGTAATCTTAACATCTGCCCTTATATGTTAAATGGT	1534
Db	2977	AGTCTC-AGTGGAGAAATCAGGTATAGTTCTGATAGGTCGCTTATATATTACTTGTCTC	3035
Qy	1535	CTTTTTCCTTGCATCTTTTAATPATCTTTCTTTGTCTATACTTTTAGTGATTTGATT	1594
Db	3036	C--TTTCCCTTACACCTTTTAATTTCTTCTTGTCTGCATATTTAGTGTTCATT	3093
Qy	1595	ATTATGCACTGTGGGAGTCTTTTCCGGTCCCATCT--ATTTGGTGTCTTGTATGCTT	1652
Db	3094	ATTATGTGGCAGGAGATTTTATTCTAGTCCAAATTTTCTTATGTTTGTGTAAGCTT	3153
Qy	1653	CTTGTACTTGTATAGGCATCTTTCTCAAGGTAGGAAATTTTCTTTTTCGTTTCT	1712
Db	3154	CTTGTACATTTAGAGGCATCTTTTTTTAAATTAGTTAGGAAATTTTAAATTAATGATTGTT	3213
Qy	1713	TGAAAAATTTTCCCTGCTTTTGTGACCTGCCCTTCT	1746
Db	3214	GGAGATGTTTCTGGGACTTTTCAGTCTCTCATCT	3247

RESULT 18

AKO50942/c

LOCUS AKO50942 2535 bp mRNA linear HTC 20-SEP-2003

DEFINITION Mus musculus 9 days embryo whole body cDNA, RIKEN full-length enriched library, clone:D030041i21 product:similar to ORF 1 [Mus musculus], full insert sequence.

ACCESSION AKO50942

VERSION AKO50942.1 GI:26341613

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

/db_xref="MG1:2418872"
 /db_xref="taxon:10090"
 /clone="D030041121"
 /tissue_type="whole body"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="9 days embryo"
 1407..1748
 /note="unnamed protein product; putative
 similar to ORF 1 [Mus musculus] (SPR|O61786, evidence:
 FASTV, 71.7%ID, 98.5%length, match=821)"
 /codon_start=1
 /protein_id="BAC34469.1"
 /db_xref="GI:26341614"
 /translation="MEIETIKKSORETLLSIENLGRSGVIDASITNRIQEKIPGP
 AEALENIDTAVKENAKCKLLQNIQEIQKMRPNLWIGIEEDSQLKPDVFNK
 IIEENPHNLKK"

CDS

Query Match 18.1%; Score 412.599; DB 1; Length 2535;
 Best Local Similarity 55.8%; Pred.No.3.8;
 Matches 851; Conservative 0; Mismatches 424; Indels 148; Gaps 11;
 798 GATCGCTATGATGATTTCTCCCAATCTCATCTGCTTAGTTTGGGTTTAAAGTCTA-TTA 856
 Db GCTGAGTGAAGTCTCTTCCCTATCTATTTTGATAACTTTGGTTGAAGTCAATTTT 2476
 857 GTCAGATATTAAATGACTGATCGGCTGCTCTTAGGCCAATTCCTTAGAATA-TCT 915
 Db ATTTGATATTAGAAATGTTATCTCAACTTGTCTTGAGACCAATTTGAITGGAAATTTG 2416
 916 TTTCCATCTTTTACTCTAAGGTGATGCTATCCATG---GTAGTTGTCTTTTGGAT 972
 Db TTTCCAGCTTTTAACTCGGAGTGTCTGTCTTGTCACTGAGTGAGTTTCTGTGAC 2356
 973 GCACAGTAGGATGATCTTTGTTTCATATCCATCTGTATACCAAGTATCTTTTCTAGA 1032
 Db ACAGCAAAATGTTGGTCCCTATTTATGATCCAGTCTGTAGTTTATTTCTTTTGGAG 2296
 1033 GAAATTAAGATCATTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1092
 Db GGAATTAG-----TCCATTTGATTTAAAGATATTAGGAAGATGATGATGATGAT 2244
 1093 CTTGTTATCTTGCACTTTGGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1152
 Db CTTGTTATCTGT 2188
 1153 GTCTCC 1212
 Db ----- 2188
 1213 CCTCTTTTGTATTTTGGCTGGAATTTATTTATTTATTTATTTTGTGATGTTGGGTAA 1272
 Db -----TAAAGATTTATTTCTGCTGTTCTAGATGATGATTTCC 2146
 1273 ATCTTTAGATGAAGTTTCTCTAGCTTCTTT--AGTCTGCAATTTGAAGATGATA 1330
 Db CTCTTTGTGTGAGTGTGCTCAATTTACTATCCCTTGAAGGCTGGAATTTGTGGAAGATA 2086
 1331 TTCTTTACATCTGATTTTATCTAGATGCTTTCTTTCTCCAACTATTCTGACAGAAAG 1390
 Db TTGTGTAATTTGTTTGTGATGGAATATCTTGGTTTCTCCATCTATGTTAAATGAGAG 2026
 1391 TTTTCTTAAGTGAAGTGTGCGCTGACATCTGTAGTCTCTTTGGAGTCTGTAGCACATC 1450
 Db TTTTACTGGGTATAGTAGCTGCTCGCATTTGTGTCTCTTAGGACGTATGACATC 1966
 1451 TGTGACGGCTCTTACATTTTGTGTTTCTTATGGAAGTCAAGTGAATTTCTAATAC 1510
 Db TGCCAGGATCTTCTAGCTTTTATGATGCTCTGTGTTGAGAAATCTGTGTAATTTCTATAG 1907
 1511 ATCTGCCCTTATATGTTAAATTTGTTCTTTTTCCTCTGCACTTTTAAATTTCTTTTCTTTG 1570
 Db GACTGCTTTAT---GTTACTTGACCTTTTTCCTTACTCTCTTTTATTTCTTTTCTTTG 1850

1571 TTCTATACCTTTTAGTGATTTGATTTATGACATGTTGGGAGTTTCTTTTCCGTFCCAAT 1630
 Db TTTTATGCAATTTGGTGTCTTGGTTATTATGTGACAGGAAGAATTTCTTTTCTGTFCCAAT 1790
 1631 CTATTGGTGTGTTGATGCTTCTGTACCTGATGAGCATCTCTTCTCAAGTTAGGA 1690
 Db CTATTGGAGTCTGTAGACCTCTTGTATATTGATGGGTATCTATTTCTTTAGGTATGG 1730
 1691 AATTTTCTTTTGGTCTTTCTGAAATATTTTCCCTGCTTTT-----GACCTGCC 1742
 Db AGTTTCTTCTATATTTTATTGAGATATCTACTGCGCCCTTTAAGTTGGATCTTCT 1670
 1743 TTCTTCCCTTCTCTATCTTCTTGGTGTGATGATGATGATGATGATGATGATGATGAT 1802
 Db TCTATACCTTATATCATAAGTTTGGTCTTCTCATTTTGTCTCGGATTTCTCTGATGAT 1610
 1803 TATGCTCGATTTATTTAGACTTAACATTTTCTTTGACCAAGGTATCCATTTCTTCTATC 1862
 Db TGGTTAGAGCTTTTACATTTTGCATTTTCTTGTGCTGTGCTGCTGCTGCTGCTGCTG 1550
 1863 TTGCTTTCACGCTGAGATTTCTTCTTCTATCTTCTGATCTGATCTGATCTGATCTGATCT 1922
 Db GCATCTTCTGACCTGGGATTTCTCTTTTATCTTCTTGTATCTTGTGTTGGTGTGATCTGCA 1490
 1923 TCTGAGGTCTCTG-TTGGGTTCTTAAATTTTCTTATTTCCAGATTTCCCTCAGTTGGGTT 1981
 Db TCTATGATCAGATCTCTTCTTCTAGTTTCTATCTCCAGGTTGTCTCCCTTTGAT 1430
 1982 TTGTTTATTAATTTTATTTTCCACTTTCAAGTCTGCTGAAATTTTACTCAATTTTCTCCCA 2041
 Db TTTCTTATTTTCTTATTTTCCATTTTAGGCTCTGGAAGGCTTTTATTTCAATCTTCAACC 1370
 2042 GATTT-----TACATTTTCAAGTTT 2063
 1369 TGTTTATGTTATTTCTTAAAGACTTCTACTGTTGACCTGTGCTGCTGCTGCTGCTGCTG 1310
 2064 CTTTAAATGATTTATTTCTTCTTCAAGGACTTTTATGAATTTCAATAAATGATGAT 2123
 Db CATTAAAGGAGTATTATGCTGTTCTTAAAGTCTCTTATGGCATCATGAGATGTAAT 1250
 2124 TAAGTCTCTTCTGCTGCTTCAAGTATGTTGCAATTTCTCAGG 2166
 Db TTAATCAAAACCTTCTTTTGGTGTGTTGGAGTAGCCAGG 1207
 RESULT 19
 BZ124878/6
 LOCUS
 DEFINITION
 CH230-490E23.TV CHORI-230 Segment 2 Rattus norvegicus genomic clone
 CH230-490E23, genomic survey sequence.
 ACCESSION
 BZ124878
 VERSION
 BZ124878.1 GI:23765825
 KEYWORDS
 GSS:
 Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE
 1 (bases 1 to 1151)
 AUTHORS
 Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,
 Shwartsbeyn,A., Gebregorgis,E., Overton,L., Russell,D., Chen,D.,
 Riggs,F., de Jong,P. and Fraser,C.M.
 TITLE
 Rat BAC End Sequences from Library CHORI-230 MboI segment
 JOURNAL
 Unpublished (1999)
 COMMENT
 Other GSSs: CH230-490E23.TJ
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org
 Clones are derived from the rat BAC library CHORI-230

(<http://www.chori.org/bacpac/rat230.htm>). For BAC library availability, please contact Pieter de Jong (pdejong@mail.chori.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/or_ering_information.htm). BAC end page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
 Plate: 490 row: E column: 23
 Seq primer: 17
 Class: BAC ends.

FEATURES
 source
 Location/Qualifiers
 1. . 1151
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /strain="BN/SSNHsd/MCW"
 /db_xref="taxon:10116"
 /clone="CH230-490E23"
 /sex="Female"
 /cell_type="Brain"
 /clone_lib="CHORI-230 Segment 2"
 /notes="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI; CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by Pieter de Jong"

Query Match 17.5%; Score 399.299; DB 1; Length 1151;
 Best Local Similarity 63.7%; Pred. No. 9;
 Matches 775; Conservative 0; Mismatches 347; Indels 95; Gaps 12;

534 TTGTGTTTGGTGAATAGTCGTAAATATC-TCTAGGTCCACTGGTTTATGACATCA 592
 Db |||||
 1146 TTGTGTTTGGTGAATAGTCGTAAATATCGGTTAAGTCCATTTGGCTCATGACTTCT 1087
 Qy |||||
 593 GTTAGTCCAGCATTTCTCGTTTGGTTTGTGAGATGACCTAACTGTTGGAGAGAA 652
 Db |||||
 1086 CTTAGTGTCTACATCTCTGTTTAACTTCTGTTT-CCAATGCTGTCATGATGAGAG 1028
 Qy |||||
 653 TGGGTTATCAAGTACGACCTACTCTGTGTGTGAGGT-CAATATGATTTTACGTGTAG 711
 Db |||||
 1027 TGGGTGGGAATCTCCCACTATTATGTGTGAGGTGCAATGTGTGTTTGGACTTTCG 968
 Qy |||||
 712 CTGTGCTGTTTATGAACTGGGTGACATTTGTTGGTGCATAGACATTAAGATTGC 771
 Db |||||
 967 TAAGTTTCTTTTACATATGATAGTGCCTTGATTTTGGGGCATAGATATTAGGATTGA 908
 Qy |||||
 772 AATGTCCTCTTGTGGGA-TTTTCCCTTGATGCCCTATGATATCTTCCCAATCTCATCT 830
 Db |||||
 907 GAGTTCATCTTGTGGGATTTTCTTTGATGAATATGAAGTGCCTTCTCTTATCTTTT 848
 Qy |||||
 831 GCTTAGTTTGGGTTTAACTCTA-TTAGTCAGATTAATAATGACTATCGCTTGCCTT 889
 Db |||||
 847 GATGACTTTTGTAGTTGAAATTTGATTTTATTGTATTTAGAAATGGCTACTCCAGCTTGC 788
 Qy |||||
 890 CTTAGGGCCATTTGCTTAGAATATC-TTTTCCATCTCTTTTACTTAAGGTGATGCTATC 948
 Db |||||
 787 CTTCTGACCATTTGTTGGAAATGTTTCCAGCCTTTTACCTGAGGTAAATGCTGTG 728
 Qy |||||
 949 CATG---GTAGGTTGCTTTTGGATGCGACAGTAGATGGATGTTGTTTCAATATCCA 1005
 Db |||||
 727 TTTGCTCTGAGGTGTGTTTCTCTAGGGAGCAGATGCGAGGTCCTCGTTGGGTATCCA 668
 Qy |||||
 1006 TTCTGTTTACCAGATCTTTTCTAGAGAAATTAAGATCATTTGATGATGTTGAGA 1065
 Db |||||
 667 GTTTGTTTAACTATGCTTTTTTATTTGGGAGTTGAG-----GCCATTTGATTTGAGA 616
 Qy |||||
 1066 ATTATCAATGACGAGTGTGTTGGATTTCTGTTATCTTGACCTTTGGAAGTGTGTGTG 1125
 Db |||||
 615 GATATTAAAGGAATAGTATTTATCTTCCCGTTTATTTTATTTTGGATGTGAGGTTATG 556
 Qy |||||
 1126 TG 1185
 Db |||||
 555 TTTGTGTG----- 548
 Qy |||||
 1186 TG 1245
 Db |||||
 547 -----CTTCACTCTCTTGTGTTGTTGTTGCCAAGATGATTAGTT 510

Qy 1246 ATTCAATATTTTCTGAATGTGGGTAAACATCTTTAGATTGAAGTTTTC--TCCTAGCCTT 1303
 Db |||||
 509 TCTTGTCTTCTTAGGATATAGCTTGCCTCCTTATGTGGCTTTTACCATTTCTTATCCT 450
 Qy |||||
 1304 CTTTAGGTCCTCAATTTGAAGATAGATATTTCTTTACATCTGATTTTATCTTAGAATGTCTT 1363
 Db |||||
 449 TTGTGCTGTGATTTGTAGAAAGATATTGTGTAATTTGGTTTGTGATGATAATCTT 390
 Qy |||||
 1364 TCTTTCTCCAACTATTTGTGACAGAAAGTTTCTTAAGTGCAGTAGTCTGGCCTGACATCT 1423
 Db |||||
 389 GGTTCCTCCGTCTATGTTAAATTGAGAGTTTTCAGAGATACAGTAACCTGGGCTGGCATTT 330
 Qy |||||
 1424 GTACTCTCTTGGAGTCTGTAGACATCTGTGACGGCTTCTTACATTTTTCAGTTTCTAT 1483
 Db |||||
 329 GTTTTCTCTTAGAGTCTGTATGACATCAGTCCACGATCTTCTGCTTCATAGTTTCT-G 271
 Qy |||||
 1484 TGGAAAGTCAGGTGTAAATCTTAATACATCTGCCCTTTATATGTTAAATGGTCTTTTTTCC 1543
 Db |||||
 270 GCGAGAAGTCTGGTGTGATTTCTGATAGTCTGCCCTTTATATGTTACTT-GACCTTTTCC 212
 Qy |||||
 1544 CTTGCACTTTTAAATATTTCTTTTCCGTCCAATCTATTGTTGTTTCTATGCTTCTTTGACCTTG 1663
 Db |||||
 211 CTTACTGCTTTTAAATATTTCTTTTCTTATTTTGTGCAITTTGGTGTCTTGTATGCT 92
 Qy |||||
 1604 TGTGGGGAGTTTCTTTTCCGTCCAATCTATTGTTGTTTCTATGCTTCTTTGACCTTG 1663
 Db |||||
 151 CGGAGAGTGTCTTTCTGCTGCCAATCTATTGAGTCTCTGAGGCTTCTTGTATGCT 92
 Qy |||||
 1664 ATAGCATCTCTTTCTCAAGGTAGGAAATTTTCTTTTGGTTTCTTGAATAATTT 1723
 Db |||||
 91 ATGGGTATCTCTTTTATAGTTAGGAAAGTTTATATATGATTTTGTGAAGATATTT 32
 Qy |||||
 1724 TCCCTGCTTTTGACCTG 1740
 Db |||||
 31 ACTGTCTTTGAGCTG 15

RESULT 20
 BZ135754
 LOCUS
 DEFINITION
 CH230-287J14.TU CHORI-230 Segment 2 Rattus norvegicus genomic clone
 BZ135754
 VERSION
 BZ135754.1 GI:23776701
 GSS.
 SOURCE
 Rattus norvegicus (Norway rat)
 ORGANISM
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE
 1 (bases 1 to 1065)
 AUTHORS
 Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K.,
 Shvartsbeyn, A., Gebregorgis, E., Overton, L., Russell, D., Chen, D.,
 Riggs, P., de Jong, P. and Fraser, C.M.
 Rat BAC End Sequences from Library CHORI-230 MboI segment
 Unpublished (1999)
 Other GSSs: CH230-287J14.TV
 COMMENT
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org
 Clones are derived from the rat BAC library CHORI-230
 (<http://www.chori.org/bacpac/rat230.htm>). For BAC library
 availability, please contact Pieter de Jong (pdejong@mail.chori.org).
 Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/or_ering_information.htm). BAC end
 page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
 Plate: 287 row: J column: 14
 Seq primer: SP6

699 ATTGTGTTCTCTTAGGGTCTGTATGACATCACTCAGGATCTTCTGGCCCTTCATAGTTT 758

1480 CTATTGGAAAAGTCAGGTGTAATCTTAATACATCTGCCTTTATATGTTAAATGGTCTTTT 1539

759 CT-GGCGAGAAGTCTGGTGTGATCTCTGATAGGTCGTGCCTTTATATGTTACTT-GACCTTT 816

1540 TTCCCTTGCATCTTTTAATATATCTTCTTCTTCTTCTATACCTTTTAGTGATTGCAATTATAT 1599

817 TTCCCTTACTGCTTTTAATATATCTTCTTCTTATTTTGTGGCTTGGTGTGTTTGCACATTAT 876

1600 GCACTGTGGGAGTCTCTTTTCGGGTGCAATCTATTTCGTGTTTCTATGCTTCTTGAC 1659

877 GTGACGGGAGTGTCTCTTTCTGTGTCCAATCTATTGAGTTCGTAGGCTTCTTGAT 936

1660 CTTGATAGGCATCTCTTCTCAAGGTTAGGAAATTTTCTTTTGGTTTTCTTGAAAT 1719

937 GCGTATGGGTATCTCTTTTATAGGTTAGGAAAGTTCTTCTATGATTTTGTGTAAGAC 996

1720 ATTTCCTGCTTTTGACCT-GCCCTTCTCCCTTCCCTATATCCCTTGGTTTTT 1773

997 ATTACTGTCTCTTGGAGCTGGAGTCTTCACTCCCTCTACCTATATTATCCCT 1051

RESULT 21

AK040990

LOCUS

DEFINITION Mus musculus adult male aorta and vein cDNA, RIKEN full-length enriched library, clone:A530056124 product:unknown EST, full insert sequence.

ACCESSION AK040990

VERSION AK040990.1 GI:26088196

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1

REFERENCE

AUTHORS Carninci, P. and Hayashizaki, Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

2

REFERENCE

AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

PUBMED 11042159

3

REFERENCE

AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, F., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE 20530913

PUBMED 11076861

4

REFERENCE

AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409, 685-690 (2001)

5

REFERENCE

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL
COMMENT

Unpublished (1999)
Other GSSs: CH230-208116.TVB
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or_ering_information.htm). BAC end
page: http://www.tigr.org/tldb/bac_ends/rat/bac_end_intro.html
Plate: 208 row: 1 column: 16
Seq primer: SP6
Class: BAC ends.

FEATURES
source

Location/Qualifiers
1..886
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SSNHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-208116"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 1"
/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by
Pieter de Jong"

Query Match 15.9%; Score 363.399; DB 1; Length 886;
Best Local Similarity 70.4%; Pred. No. 16;
Matches 564; Conservative 0; Mismatches 221; Indels 16; Gaps 6;

QY 1239 ATTATTATTCATATTTCTTGATGTTGGGTAAACATCTTTAGATGGAAGTTTTTCCTTA 1298
DB 802 ATTACTTCTGCTTTTCTAGGGTGTAGTTCTCTCTGTTGGAGTTTCCATCA 743

QY 1299 --GCCTCTTTAGTCTGCAATTTGAAGATAGATATCTTTACATCTGATTTATCTAGA 1356
DB 742 TTATCATATATAGAGCTATATATATGAAAGATATTTGATAAATTTGGTTTTCTTGTA 693

QY 1357 ATGCTTTCTTTCCCAACTATTCTGACAGAAAGTTTTCTAAGTGCAGTAGTCTGGCCT 1416
DB 582 ATATCTTTGTTTCCATCTATGTATTAATTTATAGTTTGTAGTATAGTCTGGCT 623

QY 1417 GACATCTGATCTCTTTGGAGTCTGTAGCAATCTGTGCGAGGCGCTTTTACATTTTGTAG 1476
DB 622 GCGATGTGTCTCTTAGGCTCTGTATGACATCTGTCCAGGATCTTCTGGCTTCATAG 563

QY 1477 TTTCTATTTGGAAGTCCAGGTGAATCTTAATACATCTGCTTATATATGTTTATTTGGTCT 1536
DB 562 TCGTAGT-GAGAAGTCTGGTGAATCTTAATAGCTCTGCTTTATATGTTACTT-GAGC 505

QY 1537 TTTTCCCTTGCATCTTTTAATATCTTTCTTTGTTCTATACATTTTGTAGTTGATTTAT 1596
DB 504 TTTTCCCTTACTGCTTTTAATATCTTTCTTTTATTTGTGCAATTTGGTCTTTTGAAT 445

QY 1597 TATGATGTGGGAGTTCTTTTCCGGTCCAATCTATTTTGGTTTGTATGATCTTTCTTG 1656
DB 444 TATGTTACGGGAGAGTTCTTTTCTGGTCCAATCTATTTGAAGTCTGTAGGCTTCTTG 385

QY 1657 TACCTTGATAGGATCTCTTTCTCAAGGTTAGGAATTTTCTTTTGGTTTCTTTGTAA 1716
DB 384 TATGTTATGGCATCTCTTTCTTTAGTTAGGAAGTTTCTTCTCATGATGTTTGTAA 325

QY 1717 AATATTTCCCTGCTTTTGTACCTG-----CCTTCCTCCCGCTCTCTATTCCTTTG 1767
DB 324 GATATTTACTGGCTTTTGAGCTGGGAGTCTTCACATCTCTCTATACCTATTTATCTTAG 265

QY 1768 GTTT--TTGCATAGTGTCTCTGGCTTCTCGATGCTTTTATGCTGGATTTATTTAGACTT 1825

DB 264 GTTGTATCTCATTTGAGTCTCTGATTTCTGTATGTTTATAGCCAGTAGCTTTTCTCTTT 205
QY 1826 AACATTTCTTTGACCAAGGTATCCATTTCTTATCTTCTTCTTCTTCTGCTGCTGAGATCT 1885
DB 204 TACATATCTTTGACCATTTGTCTCATGATTTCTATGGAATCTTCTGCTGAGATCT 145
QY 1886 CTCTCTATCTCTTGTATTTCTGTCTAGTGGGCTTCTCTGAGGTTCC-TGTTGGGTTCT 1944
DB 144 CTTCTCTATCTCTTGTATTTCTGTGATGCTTGTATCTATGCTCTCTCTCTCTCT 85

QY 1945 TAAATTTTTCATTTCCAGATTTCTTCTAGTTGGTTTTGTTTATTAATTTCTATTTCCAC 2004
DB 84 TTGGTTTCTTATATCCAGATTTCTCTCTTGGCTTTGCTTTCTTTATGCTTCTATTTCCAT 25

QY 2005 TTTTCAGGTCCTGAAATGTTTT 2025
DB 24 TTTTAATTCCTTCACCTGTTT 4

RESULT 23
BH348678/c

LOCUS
DEFINITION
BH348678 846 bp DNA linear GSS 03-DEC-2001
CH230-42H13.TJ CHORI-230 Segment 1 Rattus norvegicus genomic clone
CH230-42H13, genomic survey sequence.

ACCESSION
VERSION
KEYWORDSSOURCE
ORGANISM

BH348678.1 GI:17279412
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 846)
Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K.,
Shvartsbeyn, A., Gebregeorgis, B., Overton, L., Russell, D., Chen, D.,
Riggs, P., de Jong, P. and Fraser, C.M.
Rat BAC End Sequences from Library CHORI-230 EcoRI segment
Unpublished (1999)

Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org

Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or_ering_information.htm). BAC end
page: http://www.tigr.org/tldb/bac_ends/rat/bac_end_intro.html
Plate: 42 row: H column: 13
Seq primer: SP6
Class: BAC ends.

FEATURES
source

Location/Qualifiers
1..846
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SSNHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-42H13"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 1"
/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by
Pieter de Jong"

Query Match 15.9%; Score 361.399; DB 1; Length 846;
Best Local Similarity 67.0%; Pred. No. 17;
Matches 567; Conservative 0; Mismatches 261; Indels 18; Gaps 4;

QY	1210	TCCTCTTTGATTTTGGCTGGGAATATTATTATTCATATTTCTTTGAATGTGGT	1269
Db	845	TTTCTTCTTTGTTTGTGCTCCCAAGATAGTATTTTGTCTTTCTTAGGTAAGCT	786
QY	1270	AACATCTTTAGATGAAGTTTCTCTAGCCTCTTT--TAGGTCTGCAATTTGAAGATAG	1327
Db	785	GCCTCTTATGTTGGACTTTACCAATTTATATCCCTTTGTAGTCTGGATTTGTAGAAAG	726
QY	1328	ATATCTTTACATCTGATTTATCTTAGAATGCTTTCTTTCTCCCACTATTGTGACAGA	1387
Db	725	ATATTGTGTAATTTGGTTTGTGTCATGGAATATCTTTGGTTCTCCATCTATGTAATGA	666
QY	1388	AAGTTTCTTAAGTCAGTAGTCTGGCTGCACATCTGTAGTCTCTTGGAGTCTGTAGCAC	1447
Db	665	GAGTTTTCAGGATACAGTAACCTGGCTGACATTTGTGTCTTTAGGCTGTATGAC	606
QY	1448	ATCTGTGAGGGCTCTTACATTTTGAATTTCTATGTGAAAGTCAAGTGTAAATCTAA	1507
Db	605	ATCAGTCCAGGATCTTCTGGCTTCATAGTTTCT--GGCGAAAAGTCTGGTGAATCTGA	547
QY	1508	TACATCTGCTTATATGTTAAATGTTCTTTTCTCTTCCCTTGCATCTTTTAAATCTTCT	1567
Db	546	TAGTCTGCTTTATATGTTACTT--GACCTTTTCTCTTACTGCTTTTAAATCTTCTCT	488
QY	1568	TTCTTCTATCTTTTATGATTTGATTAATTAATGCACTGTGGGAGTTTCTTTCCGGTCC	1627
Db	487	TTATTTTGTGCTTTGTTGTTTGAACAATATGTCAGCAGAGTGTCTTCTTCTGCTCC	428
QY	1628	AATCTATTTGGTGTGTTGATGCTTCTGTACCTTGTATGAGGATCTTCTTCAAGGTA	1687
Db	427	AATCTATTTGGATCTGTAGGCTTCTGTATGCTGTATGCTGTATGCTTTTATAGGTA	368
QY	1688	GGAAATTTTCTTTTGTGTTTCTTGAAATATTTTCCCTGCTTTTGACCTG-----	1740
Db	367	GGGAAGTTTCTTCTATGATTTTGTGGAAGATATTTACTGGTCTTTGAGTGGGAGTCT	308
QY	1741	-----CCTTCTCCCTTCTCTATCTCTCTTTGGTTTGTGATGATGCTCTGGCTTCC	1793
Db	307	TCACCTCTCTTCTATCTATTAATCTTAGTTTGTATCTTCTCTCAATGAGTCTGATTTCC	248
QY	1794	TGATGTTTATGCTGCTGATTAATTTAGACTTAACATTTTCTTTTGACCAAGGATCCAT	1853
Db	247	TGATGTTTGGATCAGTACTTTTCCGCTTTTACATATCTTTGACAGTGTAGTCAATG	188
QY	1854	TCCTCTATCTGTCTTCACCTGCTGAGATCTCTCTCTTATCTCTTGTATCTGTCAAGTG	1913
Db	187	ATTCTATGGAATCTTCTGCTCCTGAGATCTCTCTTCCATCTCTGTATCTGTGTTG	128
QY	1914	AGGTTGCTCTGAGGTTCTGTTGGGTTCTTAAATTTTCAATTTCCAGATTTCTCTCAG	1973
Db	127	AAGCTTGTATCTACAGCTCCTTGTCTCTTTTGGTTTCTATCTCTCACTGTTGTTTCA	68
QY	1974	TTTGGGTTTGTATTAATTTCTATTTCCACTTTTCAGGCTCTGAAATGTTTACTCATTT	2033
Db	67	TGTTTCTTCTTGAATGCTCTATTTCCATTTTAACTTTTCACTGTTTGAATGTGT	8
QY	2034	TCCTCC 2039	
Db	7	TTTCCC 2	

RESULT 24
BZ221947/c
LOCUS BZ221947 825 bp DNA linear GSS 11-OCT-2002
DEFINITION CH230-361C20.TV CHORI-230 Segment 2 Rattus norvegicus genomic clone
CH230-361C20, genomic survey sequence.
ACCESSION BZ221947
VERSION BZ221947.1 GI:23880305
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;	
Rattus.	
1 (bases 1 to 825)	
Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn,A., Gebregorgis,S., Overton,L., Russell,D., Chen,D., Riggs,F., de Jong,P. and Fraser,C.M. Rat BAC End Sequences from Library CHORI-230 MboI segment Unpublished (1999)	
Other_GSSs: CH230-361C20.TV	
Contact: Shaying Zhao	
Department of Eukaryotic Genomics	
The Institute for Genomic Research	
9712 Medical Center Dr., Rockville, MD 20850, USA	
Tel: 301 838 0200	
Fax: 301 838 0208	
Email: szhao@igr.org	
Clones are derived from the rat BAC library CHORI-230 (http://www.chori.org/bacpac/rat230.htm). For BAC library availability, please contact Pieter de Jong (pdejong@mail.choi.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/or ering information.htm). BAC end page: http://www.tigr.org/tldb/bac_ends/rat/bac_end_intro.html Plate: 361 row: C column: 20 Seq primer: T7 Class: BAC ends.	
Location/Qualifiers	
1..825	
/organism="Rattus norvegicus"	
/mol_type="genomic DNA"	
/strain="BN/SsNHsd/MCW"	
/db_xref="taxon:10116"	
/clone="CH230-361C20"	
/sex="Female"	
/cell_type="Brain"	
/clone_lib="CHORI-230 Segment 2"	
/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI; CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by Pieter de Jong"	
Query Match	15.7%; Score 358.399; DB 1; Length 825;
Best Local Similarity	67.8%; Pred. No. 18;
Matches 555; Conservative	0; Mismatches 246; Indels 18; Gaps 4;
QY	1213 CCTCTTTGATTTTGGCTGGAAATATTATTATTCATATTTCTTGAATGTGGTAAC 1272
Db	821 CTCTCTCTGTTTGTGCCAAGACTATAGTCTTGTCTTTTCCAGGGTGTAGCTGC 762
QY	1273 ATCTTTAGATTGAAGTTTTTCTCCTA--GCCTCTTTTAGGTCTGCATTTGAAGATAGATA 1330
Db	761 CTTTAAATGTTGGCTTTGCCATATATTACCTTTGTAGGCTGGATTTGTAGAAAGATA 702
QY	1331 TTCTTTACATCTGATTTTATCTTAGAATGTCTTTTCTCCAACTATTTGTACAGAAAG 1390
Db	701 TTGTGTAATTTGGTTTTTGTTCATGGAATATCTGGTTTCTCCATCTATGTTAATTGAGAG 642
QY	1391 TTTTCTTAAGTGCAGTAGTCTGGCCTGACATCTGTAGTCTCTTGGAGTCTGTAGCACATC 1450
Db	641 TTTTCTGATACATGTAACCTGGCTGGCAITTTGTGTTCTCTTAGGGTCTGTATGACATC 582
QY	1451 TGTGAGGCGCTTCTTACATTTTGTAGTTTCTATTGGAAAAGTCAGGTGTAATCTTAATAC 1510
Db	581 TGTCCAGGATCTTCTGGCTTTTCATAGTCTCTGAT--GAGAAGTCTGTGTGATTTCTGATAG 523
QY	1511 ATCTGCTTTATATGTTAATTTGGTCTTTTTCCTTTGATCTTTTAAATATCTTTCTTTG 1570
Db	522 GTCTGCTTTATATGTTTACTT--GACCTTTTTCCTTTACTGCTTTTAAATATCTTTCTTTA 464
QY	1571 TTCTATACTTTTGTAGTGAATTTGATTATATATGCACTGTGGGAGTTTCTTTTCGGTCCAAAT 1630
Db	463 TTTTGTGCAATTTGGTGTGTTTGTACTATTATGTGACGGGAGGAGTTTCTTTTCTCGTCCAAAT 404
QY	1631 CTATTGGTGTGTTTGTATGCTTCTTGTACCTTGTAGGCACTCTTCTCAAGGTTAGGA 1690

[illegible]

Db 868 GTAGTGTGGATTGTAGAAAGATATTGTCTAAATTTGGTTTGTCTCATGAATATCTTGG 809
Qy 1366 TTCTCCAACTATTGTGACAGAAAGTTTTCCTAAGTCAGTAGTCTGGCCCTGACATCTGT 1425
Db 808 TTCTCCCACTAATGTTAATAGAGATTTTCTCGGATACAGTAATCTGGCTGGCAATTTGT 749
Qy 1426 AGTCTCTTGGAGTCTGTAGACATCTGTGAGGCCCTCTTACATTTTGTGATTTCTATTG 1485
Db 748 GTTCTCTTAGGGTCTGTATAACATCAGTCAGGATCTTCTGGCCCTTCATAGTTCT-GGC 690
Qy 1486 GAAAGTCAGGTGAATCTTAATACATCTGCTTTATATGTTAATTTGTTTCCCT 1545
Db 689 GAGAAGTCTGTGTGATCTGATAGGCTCCCTTATATGTTACTT-GACCTTTTCCCT 631
Qy 1546 TGCATCTTTTAAATATCTTCTTCTTGTCTTACATTTTAGTGAATTTATTTAGCACTG 1605
Db 630 TACTGCTTTTAAATATCTTCTTATTTTGTGCTTTTGGTGTGTTTGAATTTATGAGC 571
Qy 1606 TGGGAGTCTCTTTCCGGTCCAACTATTGTGTTTGTATGCTTCTTGTACCTTTGAT 1665
Db 570 GGAGGTGTTTCTTCTGTTCCAACTATTATTGGAGTCTGTAGGCTTCTGTATGCTAT 511
Qy 1666 AGGCATCTTCTTCAGAGTTAGAAATTTTCTTCTTGTGTTTCTTGAATATTTTC 1725
Db 510 GGGTATCTCTTTTATAGGTTAGGAAAGTTTCTCTATGATTTTGTGAAGATATTTAC 451
Qy 1726 CCTGCTTTTGACCTG-----CCTTCTTCCCTTCTCTATTTCTTCTTGGTT 1771
Db 450 TGGTCTTGTGAGCTGGAGTCTTCACTCTCTTATACCTATTATCTTCTTAGTTGATCT 391
Qy 1772 TTGCATAGTCTCTGGCTTCTCGATGTTTATGCTGCTGATTTATTTAGACTTAACATT 1831
Db 390 TCTCAATGAGTCTCGATTTCTCTATGTTTGGACCAAGTAGCTTTTTCGGCTTTACATT 331
Qy 1832 TTCTTTGACCAAGTATCCATTTCTCTATCTTCTTCTACTGCTGAGATTTCTCTTTC 1891
Db 330 ATCTTTGACAGTTGAGTCAATGATTTCTATGGAATCTTCTGCTCGAGATTTCTCTTC 271
Qy 1892 TATCTTGTATCTGTGAGGCTGTCTCTGAGGTTCTGTGGGTTCTTAATTTT 1951
Db 270 CATCTCTGTATTCTGTTGGTGAAGCTCGTATCACAGCTCTTGTCTCTTCTTGGTT 211
Qy 1952 TTC-ATTTCAGATTTCTTCAGTTTGGTTTGTGTTTATTAATCTATTTCCACTTCAG 2010
Db 210 TTCTATATCCAGGTTGTTTCCATGTTCTTCTCTGATTTGCTTATTTCCATTTTAA 151
Qy 2011 GTCTGAAATGTTTACTCATTTTCTCCAGTATTTACATTTTCAAGTTTCTTAAAT 2070
Db 150 TTCTTCACTGTT-----TGATTGTTTCTCGTAATTTCTTCAG 109
Qy 2071 GGATTTATTCATTTCTTCAAGGACCTTTTATGAATTCATAAAATGT 2119
Db 108 GGATTTTGGATTCCTCTCTGTAGGCTTTTACATGTTTATTAATGTTT 60

RESULT 29
BZ108872/c 920 bp DNA linear GSS 11-OCT-2002
LOCUS
DEFINITION CH230-240A15.TJ CHORI-230 Segment 1 Rattus norvegicus genomic clone
CH230-240A15, genomic survey sequence.
ACCESSION BZ108872
VERSION BZ108872.1 GI:23749751
KEYWORDS GSS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 920)
Zhao,S., Shetty,J., Shatsman,S., Teegaye,G., Geer,K.,
Sivarsbeyn,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D.,
Riggs,F., de Jong,P. and Fraser,C.M.
Rat BAC End Sequences from Library CHORI-230 EcoRI segment

JOURNAL
COMMENT

Unpublished (1999)
Other_GSSs: CH230-240A15.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or ering information.htm). BAC end
page: http://www.tigr.org/cdb/bac_ends/rat/bac_end_intro.html
Plate: 240 row: A column: 15
Seq primer: SP6
Class: BAC ends.

FEATURES
source

Location/Qualifiers
1..920
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SSHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-240A15"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 1"
/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CHORI-230 Rat (BN/SSHsd/MCW) BAC library produced by
Pieter de Jong"

Query Match 15.4%; Score 351.499; DB 1; Length 920;
Best Local Similarity 67.6%; Pred. No. 17;
Matches 561; Conservative 0; Mismatches 250; Indels 19; Gaps 5;
Qy 1213 CTTCTTTGATTTTGGCTGGAAATATTATATATCATATTTCTTGAATCTGGGTAAAC 1272
Db 844 CTTCTCTTTGTTGTGCCCCAGTGAATTAGTTTCTGCTTTCTAAGGGGTAGCTGTC 785
Qy 1273 ATCTTTAGATTCAGATTTTCTCCAGCTTCTT--TAGGCTGCTATTTGAGATAGATA 1330
Db 784 CTCCTTATGTTGGGCTTTACCAATTTATTATCTTTGTAGTCTGGAATTTAGGAAGATA 725
Qy 1331 TTCTTTACATCTGATTTTATCTTAGAATGCTTTCTTTCTCCAACTATTGTGACAGAAAG 1390
Db 724 TTATGTAATTTGGTTTGTGATGAATATCTTGTCTTCTCCATCTATGTTAATTTAGAG 665
Qy 1391 TTTTCTAAGTCAGTAGTCTGCTGCTGACATCTGATCTCTTGGAGTCTGTAGCACATC 1450
Db 664 TTTTCTGGATACAGTAACCTGGCTGGCAATTTGTGTTCTCTTAGGGTCTGTATACATC 605
Qy 1451 TGTGAGGGCCCTCTTACATTTTGTAGTTTCTATTGAAAAGTCAGGTGTAAATCTTAATAC 1510
Db 604 TGTCCAGGATCTTCTGGCTTTCATATCTCCGGT-GAGAAATCTGCTGATTTCTGATAG 546
Qy 1511 ATCTGCTTTATATGTTAATTTGGTCTTTTTCCTTGTGATCTTTTAAATATTTCTTTG 1570
Db 545 GTCTGCTTTATATGTTTACTT-GACCAATTTCCGTTACTGCTTTTAAATATTTCTTCTTTG 487
Qy 1571 TTCTATATCTTTAGTATTTGATTTATTTATGCACTGTGGGAGTTTCTTTTCCGGTCCAAAT 1630
Db 486 TTTTGTGCAATTTGGTGTGTTTGACTATTATGACAGGAGAGTTTCTTTTCTGGCCCAAT 427
Qy 1631 CTATTGTGTTTGTATGCTTCTTTGTACCTTTAGGATCTCTTTCTCAAGGTTAGGA 1690
Db 426 ATATTGGAGTTCTGTATGCTTCTTGTATGTTTATGAGGATCTCTTTCTTTAGGTTAGG 367
Qy 1691 AATTTTCTTTTGGTTTCTTGAATAATTTTCCCTGCTTCTTGAACCTG----- 1740
Db 366 AAGTTTCTCTATGATTTTGTGAAGATATTTACTGTCTCTTTGAGCTGGAGGCTTCA 307
Qy 1741 ----CCTTCTTCCCTTCTCTCTATTCTTGTGTTTGTGATAGTCTCTCTGCTTCTCTGG 1796

Db 306 CTCCTCTTCTACACCTGTTATCTTAGGTTTCACTCTCTCATTGAGTCTCGGATTTCTCTGT 247
 Qy 1797 ATGTTTATGCTGGATATTTAGACTTAACATTTCTTTGACCAAGGTATCAATTTCT 1856
 Db 246 ATGTTTGGACAGTAGTTTTTCCATTTTCCATTTCTTTGACTGTGTGTCATGATTT 187
 Qy 1857 TCTATCTGCTTCCACCTGCTGAGATCTCTCTTATCTCTGTAATCTGTCAGTGAGG 1916
 Db 186 TCTATGAACTTCTGCTGCTGAGATTTCTCTTATCTCTGTAATCTCTTTGATG 127
 Qy 1917 CTGCTCTGAGGTTCCTGTTGGTCTTAATTTTTCATTTTCCAGATTTCTTCAGTT 1975
 Db 126 CTGCTATCTGCTGCTCTGAGATTTCTCTTATCTCTGTAATCTCTTTGATG 127
 Qy 1976 TGGCTTTTGTATTAATCTTATTTCCATTTTCCAGTCTGCTGAAATGTTT 2025
 Db 66 TGTGCTTTCTTATGCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 17

RESULT 30
 BZ109050/c
 LOCUS
 DEFINITION CH230-218C4.TJ CHORI-230 Segment 1 Rattus norvegicus genomic clone
 CH230-218C4, genomic survey sequence.
 ACCESSION BZ109050
 VERSION BZ109050.1 GI:23749929
 KEYWORDS GSS.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 823)
 AUTHORS Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K.,
 Shvartsbeyn, A., Gebregorgis, E., Overton, L., Russell, D., Chen, D.,
 Riggs, P. de Jong, P. and Fraser, C.M.
 Rat BAC End Sequences from Library CHORI-230 EcoRI segment
 JOURNAL Unpublished (1999)
 COMMENT Other GSSs: CH230-218C4.TJ
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org
 Clones are derived from the rat BAC library CHORI-230
 (<http://www.chori.org/bacpac/rat230.htm>). For BAC library
 availability, please contact Pieter de Jong (pdejong@mail.cho.org).
 Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/orering_information.htm). BAC end
 page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
 Plate: 218 row: C column: 4
 Seg primer: SP6
 Class: BAC ends.

FEATURES
 Location/Qualifiers
 1..823
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /strain="BN/SENHsd/MCW"
 /db_xref="taxon:10116"
 /clone="CH230-218C4"
 /sex="Female"
 /cell_type="Brain"
 /clone_lib="CHORI-230 Segment 1"
 /note="Vector: pTARAC2.1; Site 1: EcoRI; Site 2: EcoRI;
 CHORI-230 Rat (BN/SENHsd/MCW) BAC library produced by
 Pieter de Jong"
 Query Match 15.4%; Score 351.099; DB 1; Length 823;
 Best Local Similarity 67.3%; Pred. No. 19;
 Matches 561; Conservative 0; Mismatches 239; Indels 33; Gaps 5;

Qy 1301 CTTCTTTAGGCTGCTGATTTGAAGATAGATATTTCTTTACATCTGATTTTATCTTGAATGT 1360
 Db 819 CTTTGTAGTCTGCTGATTTGTAGAAAGATAATTGTAAATTTGGTTTGTTCATGGAATAT 760
 Qy 1361 CTTTCTTTCTCCAACTATTTGACAGAAAGTTTTCCTAAGTGCAGTAGTCTGGCTGACA 1420
 Db 759 CTTGTTTTCTCCATCAATTAATTGAGAGTTTGTCTGGAATACACCTGGGCTGGCA 700
 Qy 1421 TCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGAGGGCTTCTTACATTTTTCAGTTTC 1480
 Db 699 TTTGTGTTCTTTAGGCTCTGTATGACATCAGTCCAGGATATTCGGGCTTCATAGTTTC 640
 Qy 1481 TATTGGAAAGTCAAGTGTAAATCTTAATACATCTGCGCTTTATATGTTAATTTGGTCTTTT 1540
 Db 639 T-GGCGAGAAGTCTGCTGATTTCTGATAGGTCTCCCTTTATATGTTACTT-GACCTTTT 582
 Qy 1541 TCCCTTGCATCTTTTAAATATTTCTTTTGTCTATATCTTTTAGTGATTTGATTATATG 1600
 Db 581 TCCCTTACTGCTTTTAAATATTTCTTTTATTTTGTGCTTTGGTTTGGACATATG 522
 Qy 1601 CACTGTGGGAGTTTCTTTTCCGTCCTTCTTCCCTTCTCTTATTTCC-----TTTG 1767
 Db 401 TTTACTGCTCTTTGAGCTGGAGTCTTCACTCTCTTCTATACCTATTATCTTAGTTTG 342
 Qy 1768 GTTTTGCATAGTGTCTCTGCTTCTCGGATTTTATGCTTGGATTTATTTAGACITAA 1827
 Db 341 ATCTTCTCATTTGAGTCTGCTGATTTCTGTATGTTTGGACAGTAGTCTTCCACITTA 282
 Qy 1828 CATTCTTTTGGACCAAGGTATCCATTTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCT 1887
 Db 281 CATTATCTTTGACAGTTGAGTCAATGATTTCTATGGAATCTTCTGCTCTGAGATCTCT 222
 Qy 1888 CTTCTATCTTTGATTTCTGTCAGTGAGGTGCTCTGAGGTCTCTGAGGTCTCTTGGTTCTTAA 1947
 Db 221 CTTCCATCTCTGATTTCTGTTGGTGAAGTTGTATCTACAGCTCTTGTCTCTCTTTT 162
 Qy 1948 TTTTCTT-ATTTCAGATTTCTTCTGAGTTTGGTTTGTGTTTATTAATTTTCTTCTTCT 2006
 Db 161 GGTCTTCTATATCCAGGTTGTTTCCATGTTCTTTCTTCTGATGCTTCTATTTCCATT 102
 Qy 2007 TCAGTCTCTGAAATGTTTACTCAATTTCTCCAGTATTTACATTTTCAAGTTTCTT 2066
 Db 101 TTAATTTCTTCAACTGTTT-----GATTGTTTCTTCTTCTTCTTCTTCTTCTT 59
 Qy 2067 TAAATGATTTTATTTCAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2119
 Db 58 TCAGGATTTTGTGCT 6

RESULT 31
 BZ184947/c
 LOCUS
 DEFINITION CH230-329122.TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone
 CH230-329122, genomic survey sequence.
 ACCESSION BZ184947
 VERSION BZ184947.1 GI:23835048
 KEYWORDS GSS.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 816)
 AUTHORS Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K.,

Shvartsbeyn, A., Gebregorgis, E., Overton, L., Russell, D., Chen, D.,
 Riggs, P., de Jong, P. and Fraser, C.M.
 Rat BAC End Sequences from Library CHORI-230 MboI segment
 Unpublished (1999)
 Other GSSs: CH230-329122.TV
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org
 Clones are derived from the rat BAC library CHORI-230
 (<http://www.chori.org/bacpac/rat230.htm>). For BAC library
 availability, please contact Pieter de Jong (pdejong@mail.cho.org).
 Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/orering_information.htm). BAC end
 page: http://www.tigr.org/tadb/bac_ends/rat/bac_end_intro.html
 Plate: 329 row: 1 column: 22
 Seq primer: SP6
 Class: BAC ends.

FEATURES
 source
 1..816
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /strain="BN/SSNHsd/MCW"
 /db_xref="taxon:10116"
 /clone="CH230-329122"
 /sex="Female"
 /cell_type="Brain"
 /clone_lib="CHORI-230 Segment 2"
 /notes="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
 CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by
 Pieter de Jong"

Query Match 15.4%; Score 349.899; DB 1; Length 816;
 Best Local Similarity 69.7%; Pred. No. 19;
 Matches 525; Conservative 0; Mismatches 211; Indels 17; Gaps 4;

QY 1297 TAGCCTTTTGTAGTCTGCAATTTGAAGATAGATATTTTACATCTGATTTATCTTAGA 1356
 DB 782 TAATCCCTTTGTAGTCTGGAGTTGGAGAAAGATATTGTGACATTTGGTTTGTCTATGGA 723
 QY 1357 ATGCTTTCTTTCTCCAACTATTGTGACAGAAAGTTTTCATAGTCAGTAGCTGGCCT 1416
 DB 722 ATTTCTTGGTTTTCCTCATCTATGTTAATGAGAGTTTTCAGGATACAGTAACCTGGCT 663
 QY 1417 GACATCTGATCTCTTTGGAGTCTGTAGCATCTGTGACATCTGTGAGGCGCTTCTTACATTTTGA 1476
 DB 662 GGCATTTGTGTTCTTTAGGCTGTGTATGACATCTGTCCAGGATCTCTGGCTTTCTAG 603
 QY 1477 TTTCTATTGGAAGAGTCAGGTGTAATTTCTAATACATCTGCTTTATATGTAATTTGGTCT 1536
 DB 602 TTTCTGCT-GAGAACTCTGGTGAATTTCTTATAGTCTGCTTTTATGTTACTTGATC- 545
 QY 1537 TTTTCTCCCTGCATCTTTTAAATCTTTCTTTGTTCTTACATTTTGTAGTATTTGATTAT 1596
 DB 544 TTTTCTCCCTACTACTTTTAAATCTTTCTTTGTTGTGCAATTTGGGTTTGTACTAT 485
 QY 1597 TATGCACTGTGGGAGTTTCTTTTCGGTCCAACTATTATTTGGTGTGTTTGTATCTTTG 1656
 DB 484 TATGTGACAGAGAGTTTCTTTCTGGTCCAACTATTATTTGGAGTTCTGTAGGCTTTCTG 425
 QY 1657 TACCTTGATAGGATCTTTCTCAAGTTAGGAATTTTCTTTTGTGTTTCTTTGTTTCTTGA 1716
 DB 424 TATGTTTATGGGCACTCTTTCTTTAGGTTAGGAAGTTTCTTCTTACAAATTTTGTGAA 365
 QY 1717 AATATTTTCCCTGCTTTTGACCTG-----CCTTCTTCCCTTCTCTATTC 1762
 DB 364 GATATTTTCTGGTCTTTGAGCTGGAGTCTTCACTCTCTTCTATACAGTATCTCTTAG 305
 QY 1763 CTTTGGTTTTGCAATGATGCTCTGGCTTCTGAGATGTTTATGCTGGATTAATTTTGA 1822

DB 304 GTTGACCCCTCATTTGTGCTCTGATTTTCTGATGTTTGGGCACTACATTTTCTG 245
 QY 1823 CTTACATTTTCTTTGACCAAGGTATCCATTTCTTCTATCTTGTCTTCTACCTGCTGAGAT 1882
 DB 244 TTTTACATTTCTTCATCAGTTGTGCGATGATTTCTTAAGCAATCTTCGGCTCTGAGAT 185
 QY 1883 TCTCTCTTCTATCTCTTGTATTTCTGTCAGTGAGGCTTCTCTGAGGTTC-TGTGGT 1941
 DB 184 TCTCTCTTCTATCTCTTGTATTTCTGTTGTTGATGCTTGTATCTGCGGCTCTTGTCTCT 125
 QY 1942 TCTTAATTTTTCATTTCCAGATTTCTTTCAGTTGGTGTGTTTGTATTAATTTCTATTC 2001
 DB 124 CTTTGGTTTCTATATCCAGGTTTGTTCCTTTGTGCTTTCTTATTACTTCTATTTTC 65
 QY 2002 CACTTTCAGGCTCGAAATGTTTACTCATTTT 2034
 DB 64 CATTTTCAATTCCTTCACCTGTTTGATTTGTT 32

RESULT 32
 BZ097017/c
 LOCUS
 DEFINITION
 CH230-22306.TJB CHORI-230 Segment 1 Rattus norvegicus genomic clone
 CH230-22306, genomic survey sequence.

ACCESSION
 BZ097017
 VERSION
 BZ097017.1
 GI:23737901
 KEYWORDS
 GSS.
 SOURCE
 Rattus norvegicus (Norway rat)

ORGANISM
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE
 1 (bases 1 to 910)
 Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K.,
 Shvartsbeyn, A., Gebregorgis, E., Overton, L., Russell, D., Chen, D.,
 Riggs, P., de Jong, P. and Fraser, C.M.
 Rat BAC End Sequences from Library CHORI-230 EcoRI segment
 Unpublished (1999)
 Other GSSs: CH230-22306.TVB

COMMENT
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org
 Clones are derived from the rat BAC library CHORI-230
 (<http://www.chori.org/bacpac/rat230.htm>). For BAC library
 availability, please contact Pieter de Jong (pdejong@mail.cho.org).
 Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/orering_information.htm). BAC end
 page: http://www.tigr.org/tadb/bac_ends/rat/bac_end_intro.html
 Plate: 223 row: 0 column: 6
 Seq primer: SP6
 Class: BAC ends.

FEATURES
 Location/Qualifiers
 1..910
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /strain="BN/SSNHsd/MCW"
 /db_xref="taxon:10116"
 /clone="CH230-22306"
 /sex="Female"
 /cell_type="Brain"
 /clone_lib="CHORI-230 Segment 1"
 /note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
 CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by
 Pieter de Jong"

Query Match 15.3%; Score 349.699; DB 1; Length 910;
 Best Local Similarity 67.3%; Pred. No. 17;
 Matches 561; Conservative 0; Mismatches 253; Indels 19; Gaps 5;

ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 878)
 AUTHORS Zhao, S., Shetty, J., Shatsman, S., Teegaye, G., Geer, K.,
 Shvartsbeyn, A., Gebregorgis, E., Overton, L., Russell, D., Chen, D.,
 Riggs, F., de Jong, P. and Fraser, C.M.
 TITLE Rat BAC End Sequences from Library CHORI-230 EcoRI segment
 JOURNAL Unpublished (1999)
 COMMENT Other GSSs: CH230-238N16.TJ
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org
 Clones are derived from the rat BAC library CHORI-230
 (http://www.chori.org/bacpac/rat230.htm). For BAC library
 availability, please contact Pieter de Jong (pdejong@mail.cho.org).
 Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/or ering_information.htm). BAC end
 page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
 Plate: 238 row: N column: 16
 Seq primer: T7
 Class: BAC ends.

FEATURES
 source Location/Qualifiers
 1..878
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /strain="BN/SsNHsd/MCW"
 /db_xref="taxon:10116"
 /clone="CH230-238N16"
 /sex="Female"
 /cell_type="Brain"
 /clone_lib="CHORI-230 Segment 1"
 /notes="vector: pFAR2.1; Site_1: EcoRI; Site_2: EcoRI;
 CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
 Pieter de Jong"

Query Match 15.3%; Score 348.299; DB 1; Length 878;
 Best Local Similarity 67.6%; Pred No 18;
 Matches 556; Conservative 0; Mismatches 247; Indels 19; Gaps 5;

Qy 1215 TCCTTTGATTTGGCTGGAATATTATATATCATATTTCTTGAATGCGGTAAACAT 1274
 |||||
 Db 821 TCCTCTGTTGGTGGCCCAAAATGAATTAATTCCTGTTCTTTCTAGGGTATAACTTGCCT 762

Qy 1275 CTTTAGATTCAGATTTTCTCCAGCTTCTT--TAGGCTGCATTTGAAGATAGATATT 1332
 |||||
 Db 761 CTTATCTTGGGCTTTACCATATATATATCTCCAGCGCTAGATTGTAGAAAGATATT 702

Qy 1333 CTTTACATCTGATTTATCTTAGAATGCTCTTCTTCTCCAACTATTGTGACAGAAAGTT 1392
 |||||
 Db 701 GTGTAAATTTGGTTTGTGATGAATATCTTGGTTCTCCATCTATGTTAACTGAGATT 642

Qy 1393 TTCTAAGTCAGTAGTCTGGCTGACATCTGATGCTCTTGTGAGTCTGTAGCAGATCTG 1452
 |||||
 Db 641 TTGAGGATACAGTAACCTGGCTGGCAATTTGTCTCTTAGGGTCTGTATGACATCAG 582

Qy 1453 TGACAGGCCCTTCTACATTTTGATTTCTATTGGAAGAGTCAGTCTTAATTTCAATACAT 1512
 |||||
 Db 581 TCAGGATCTCTGGCTTTCATAGTTCT--GGAGAGAGCTGTGTTGATTCGTAGCT 523

Qy 1513 CTGCTTTATATGTTAATTCGGCTTTTTCCTTCGATCTTTTAAATCTTTTCTTTGTT 1572
 |||||
 Db 522 CTGCTTTATATGTTACTT--GACCTTTTCCCTTACTGCTTTTAAATCTTTCTTTTAT 464

Qy 1573 CTATACCTTTTGTGATTTGATTTAATGCACTGTGGGAGTTTCTTTCCGGTCCAATCT 1632
 |||||
 Db 463 TTGTGATTTGGTGTGTTTGTACTATTATGACGGGAGGTGTTCTTTCTGTTCCTCAATCT 404

Qy 1633 ATTTGGTGTGTTTGTATGCTTCTTGACCTTGATAGGCATCTCTTTCTCAAGTTAGGAAA 1692
 |||||
 Db 403 ATTTGGAGTTCGTAGAGGTTCTTGATGCTTATGGGTAATCTCTTTTATTAGTTAGGAA 344

Qy 1693 TTTTCTTTTGGTGTGTTTCTTGAATAATTTTCCCTGCTTTTGACCTG----- 1740
 |||||
 Db 343 GTTTCTTCTATGATTTTGTGAGATATTTACTGGTCTCTTGAGCTGGGAGTCTTCACT 284

Qy 1741 --CCTTCTTCCCTTCCCTTATCTCTTGGTGTGTTTGTGATAGTCTCTGGGCTCCTGGAT 1798
 |||||
 Db 283 CTCCTTCTATACCTATATCTTAGGTTTGAATCTTCTCAATGAGTCTCGATTTCTGTAT 224

Qy 1799 GTTTTATGCTCGGATTTATTTAGACTTAACATTTCTTTTGACCAAGTATCCATTTCTTC 1858
 |||||
 Db 223 GTTTTGGACCAATTAGCTTTTTCGCTCTACATTTTGTGACAGTGAATCAATGATTTTC 164

Qy 1859 TATCTGTCTTCACTGCTGAGATCTCTCTCTCTATCTCTTGTATCTGTGATCTGAGTGGCT 1918
 |||||
 Db 163 TATTTAATCTTCTGCTCTGAGATCTCTCTCTCTATCTCTTGTATCTTGTGTTGGTGAAGCT 104

Qy 1919 TGCTCTCAGGTTCTCTGTTGGTCTTAAATTTTTC-ATTTCCAGATTTCTTCTCAGTTTG 1977
 |||||
 Db 103 CGTATCTACAGCTCTGCTCTCTCTCTCTGTTTCTATATCCAGGGTGTGTTTCCAATGTG 44

Qy 1978 GGTGTTGTTTATTAATCTTATTTCCACTTTTCAGTCTCGTGA 2019
 |||||
 Db 43 TTCTTTCTTGATGCTTCTTATTCATTTTAAATTCCTTCAA 2

RESULT 36
 AK088420/c 2559 bp mRNA linear HTC 20-SEP-2003
 LOCUS AK088420
 DEFINITION Mus musculus 2 days neonate thymus cells cDNA, RIKEN
 full-length enriched library, clone:E430016D20 product:similar to
 PORF1 [Mus musculus domesticus], full insert sequence.

ACCESSION AK088420
 VERSION AK088420.1 GI:26104798
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636

REFERENCE 2
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159

REFERENCE 3
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kishimoto, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861

REFERENCE 4
 AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL
REFERENCE
AUTHORS

Nature 409, 685-690 (2001)

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL
REFERENCE
AUTHORS

Nature 420, 563-573 (2002)

6 (bases 1 to 2559)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscl.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers

FEATURES
source

1. 2559
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="FANTOM DB:E430016D20"
/db_xref="MG1:2427317"
/db_xref="taxon:10090"
/clone="E430016D20"
/cell_type="thymic cells"
/tissue_type="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="2 days neonate"
/notes="similar to PORF1 [Mus musculus domesticus] (SPTRIQ91V68, evidence: FASTV, 70.8%ID, 100%length, match=1094)"]

misc_feature

1. 2559 Score 347.299; DB 1; Length 2559;
Best Local Similarity 66.28; Pred. No. 6.8;
Matches 584; Conservative 0; Mismatches 277; Indels 21; Gaps 6;

Query Match
Best Local Similarity 66.28; Pred. No. 6.8;
Matches 584; Conservative 0; Mismatches 277; Indels 21; Gaps 6;

1232-TGGAATTATTATTATTCATPATTTCCTGTAATGCGGTGAACATCTTTAGATTGAAAGTTT 1291
1495 TTGAAGATTACTTCTTGTCTTTCAAGGCTAGTTCCTTATAGCTGTTTC 1436
1292 TCTCCAGCTTTTGGTTCGATTTGAAGATAGATATCTTACATCTGATTTTAC 1351
1435 ATCTATTATCTTTGTAGGCTCGATTTGTAAGATATTTGTAATTTGGATTGTC 1376
1352 TTGAATGTCTTCTTCTCCAACTATTGTGCAGAAAGTTTCTTAAGTCAGTAGTCT 1411
1375 ATGGAATACCTTGGTTTCTCCGCTATGTTAACTGAGAGTTTGTGATGGGTATAGTACC 1316

1412 GGCTGACATCTGTAGTCTCTTTGGAGTCTGTGACATCTGTGAGGCTTCTTACATT 1471
1315 CACCTGGCATTTGTGTCTCTTAGGGTATGTATGACATCTGCCAGGATCTTCTAGCTT 1256
1472 TTGAGTTTCTATTGSAAGTCAAGTGTAAATCTTAATACATCTCCCTTTATATGTTAAT 1531
1255 CATGCTCTCTGGT-GAGAAGTCTGTGTAAATCTGAAGGCTGCCCTTATATGTTACTT 1197
1532 GGTCTTTTCCCTTGGCATCTTTTAAATATCTTTCTTTGTGTCTATACCTTTTGTAGTTG 1591
1196 -GACCTTTTCCCTTACCATTTAAATATCTTTCTTTGTGTAGTGCATTTGGTGTTCG 1138
1592 ATTATTATGCACTGTGGGAGTCTTTTCCGGTCCATCTATTGTGTGTGTGTGTGTGCT 1651
1137 ATTATTATGATGGGGAATTTCTTTTCTGGTCCAGTCTATTAGAGTTCTGTAGCT 1078
1652 TCTTGTACCTTGTATAGGATCTCTTTTCTCAAGGTTAGGAATTTTCTTTTGTGTTTC 1711
1077 TCTTGTATGTTATGCGGATCTCTTTCTTAGTGTAGGGAGTTTCTTCTATATTTTG 1018
1712 TTGAATATTTTCCCTGCTTTTGCCTG-----CCTTCTTCCCTTCTCTC 1757
1017 TTGAAGATATTTTACTTGGCCCTTTAAGTTGGGAATCTTCACTCTCTTATACCTATTATC 958
1758 TATCTCTTGTGTTTGTGATAGTCTCTGCTTCCCTGGATGTTTATGCTGATTAAT 1817
957 CTAGTTTCACTCTTCTTGTCTGCTGATTTTCTGGAATTTGGGTAGATCTTT 898
1818 TTAGATTTAAATTTCTTTTGAACCAAGTATCCATTTCTTCTTGTCTTCTCACTGCT 1877
897 TTGCTTTTGTCAATTTCTTGTGCTGCTGATTTTCTGGAATTTCTTCTTCTTCACT 839
1878 GAGATCTCTCTCTCTCTCTCTGTTGATTTGTCAGTGGCTGCTCTGAGGTTCCCTG-T 1936
838 GAGATCTCTCTCTCTCTCTCTGTTGATTTTCTGTTGATGCTGCACTGATGACTCCGAT 779
1937 TGGGTCTTAAATTTTCTTCAATTTCCAGATTTCTTCAAGTTTGGGTTTGTGTTTATTAATCT 1996
778 TCTTCTCTAGTTTCTTCAACCCAGGTTCTCTCCCTTGTGATTTCTTATTTGTTCT 719
1997 ATTCCACTTTTCAAGTCTGTAATGTTTCTTCACTTCTTCTTCTTCCAGTAT---TTACATTT 2053
718 ATTCCATTTTGTAGATCTGATGTTTGTGTTTCTTCACTTCTTCTTCACTTTTGTGATTTGCTT 659
2054 TCATAGGTTTCTTAAATGCAATTTATTTCAATTTCTTCTTCTTCAAGG 2095
658 TCTGTATTTCTTAAAGGATTTGTTGTTTCTTCTTCTTAAAG 617

BZ097284 835 bp DNA linear GSS 10-OCT-2002
CH230-236115-TV CHORI-230 Segment 1 Rattus norvegicus genomic clone
CH230-236115, genomic survey sequence.
BZ097284
BZ097284.1 GI:23738168
GSS.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 835)
Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K.,
Shvartsbeyn, A., Gebregorgis, E., Overton, L., Russell, D., Chen, D.,
Riggs, F., de Jong, P. and Fraser, C.M.
Rat BAC End Sequences from Library CHORI-230 EcoRI segment
Unpublished (1999)
Other GSSs: CH230-236115.TU
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research

RESULT 37
BZ097284/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the rat BAC library CHORI-230

(<http://www.chori.org/bacpac/rat230.htm>). For BAC library

availability, please contact Pieter de Jong (pdjong@mail.cho.org).

Clones may be purchased from BACPAC Resources

(http://www.chori.org/bacpac/orering_information.htm). BAC end

plate: http://www.tigr.org/tldb/bac_ends/rat/bac_end_intro.html

Plate: 236 row: I column: 15

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

1..835

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/strain="BN/SENHsd/MCW"

/db_xref="taxon:10116"

/clone="CH230-236I15"

/sex="Female"

/cell_type="Brain"

/clone_lib="CHORI-230 Segment 1"

/notes="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;

CHORI-230 Rat (BN/SENHsd/MCW) BAC library produced by

Pieter de Jong"

Query Match

Best Local Similarity 67.1%; Pred. No. 19; Length 835;

Mismatches 559; Conservative 0; Mismatches 255; Indels 19; Gaps 5;

QY 1221 GATTTTGGCGCTGGAATATTTATTTATTTCTTGAATGCGGTAACATCTTTAG 1280

835 GTTTTGGCCCCAAGACGATAGTTTCTGCTTCTTAGGGTATAGCTTGCCCTCTAT 776

QY 1281 ATTGAAGTTTTTCTCTAGCCTTCTT--TAGGTCTGCAITTTGAAGATAGATATTTCTTTAC 1338

775 GTTGGGCTTTACCATTATTTATCTTTGTAGTCTGGAITTTGAGAAGATATTTGTTAA 716

QY 1339 ATCTGATTTATCTTAGAATGCTTTCTTCTTCCAACTATTTGTGACAGAAAGTTTCTTA 1398

715 ATTGTGTTTGTGATGGAATATCTTGGTTTCTCCATCAATGTTAATGTAGAGTTTGTCTG 656

QY 1399 AGTGCAGTACTGGCTGACATCTGAGTCTCTTGGAGTCTGTAGACATCTGTGCAGG 1458

655 GATACAGTAACTGGGTGCGATTTGTGTTCTTTAGGTCTGTATAAATGATGCTCCAG 596

QY 1459 GCCTTCTCATTTTGTAGTTTCTATTGGAAAGTCAAGGTGTAATTTCTAATACATCTGCT 1518

595 ATCTTCTGGCCTTCATAGTTTCT--GGCGAAGTCTGGTGTGATTTCTGATAGTCTCCCT 537

QY 1519 TTATATGTTAATTTGGTCTTTTCCCTTGCATCTTTATATATTTCTTTCTGTCATAC 1578

536 TTATATGTTAATTTGGTCTTTTCCCTTGCATCTTTATATATTTCTTTCTGTCATAC 478

QY 1579 TTTTAGTGATTTGATTTATGCACTGTGGGAGTTTCTTTTCCGTCCTCAATCTATTGG 1638

477 GTTTGGTGTGTTTGACAAATATGTGACGGGAGGTGTTCTTTTCTGTCCTCAATCTATTGG 418

QY 1639 TGTGTTGTATGCTTCTGTACCTGTAGGATCTTTCTTCAAGTTAGGAATTTTTC 1698

417 AGTCTGTAGGCTTCTGTATGTCTATGGGTATCTTTTTPAGTTAGGGAAGTTTTC 358

QY 1699 TTTTGTGTTTCTTGAAAATATTTTCCCTGCTTTTGAACCTG-----CCTT 1744

357 TTCTATGATTTTCTTGAAGATATTTACTGCTCTTTGAGCTGGAGTCTTCACTCTCTTC 298

QY 1745 CTTCCCTCTCTATCTCTTGGTTTGTGATGATGCTCTGCGCTCTCTGGAATTTTTC 1804

297 TATACCTATTATCTTAGGTTTGTATCTCTCATGTGTCTGCTGATTTCTTGTATGTTTG 238

QY 1805 TGCTGTGATTTTATAGACTTATACATTTTCTTTTGACCAAGGTATCCATTTCTTCTATCTT 1864

Db 237 GACCAGTAGCTTTTTCGCTTTTACATTATCTTTGACAGTTGAGTCAATGATTTCTATGGA 178

QY 1865 GTCTTCACTGCTGAGATTTCTCTTCTATCTCTTTGATTTCTGTCAGTGAGGCTTGTCTC 1924

Db 177 ATCTTCTGCTCTGAGATTTCTCTCTTCAATCTCTTTGATTTCTGTTGAGCTTGTATC 118

QY 1925 TGAGTTTCTGTTGGGTTCTTAAATTTTTC-ATTCCAGATTTCTTCCAGTTTGGGTTTT 1983

Db 117 TACAGCTCTTGTCTCTCTTCTTTGTTTCTATATCCAGGCTTGTTCCTCATGTTCTTTT 58

QY 1984 GTTATTTAATTTCTATTTCCACTTTTCCAGTCTGCTGAAATGTTTACTCATTTTCC 2036

Db 57 CTGATTCCTCTATTTCCATTTTAACTCTTCACTGTTGATGTTTTC 5

RESULT 38

BZ106482/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Rattus norvegicus (Norway rat)

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 833)

REFERENCE

AUTHORS

SHVARTSBEYN, A., GEORGE, G., OVERTON, L., RUSSELL, D., CHEN, D.,

RIGGS, F., DE JONG, P. and FRASER, C.M.

Rat BAC End Sequences from Library CHORI-230 EcoRI segment

Unpublished (1999)

Other GSSs: CH230-228A7.TJ

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the rat BAC library CHORI-230

(<http://www.chori.org/bacpac/rat230.htm>). For BAC library

availability, please contact Pieter de Jong (pdjong@mail.cho.org).

Clones may be purchased from BACPAC Resources

(http://www.chori.org/bacpac/orering_information.htm). BAC end

plate: http://www.tigr.org/tldb/bac_ends/rat/bac_end_intro.html

Plate: 228 row: A column: 7

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

1..833

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/strain="BN/SENHsd/MCW"

/db_xref="taxon:10116"

/clone="CH230-228A7"

/sex="Female"

/cell_type="Brain"

/clone_lib="CHORI-230 Segment 1"

/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;

CHORI-230 Rat (BN/SENHsd/MCW) BAC library produced by

Pieter de Jong"

Query Match

Best Local Similarity 67.3%; Pred. No. 19;

Mismatches 557; Conservative 0; Mismatches 252; Indels 19; Gaps 5;

QY 1209 CTCCTCTTTTGTATTTTGGCTTGGGAATTTATTTATTTATTTTCTTGAATGGG 1268

Db 827 CTTCCATCTCTTCTTTTGTGCCCCAAGACGATTTCTGCTTCTTCTAGGTATAGC 768


```

REFERENCE
AUTHORS      1. (bases 1 to 838)
              Zhao,S., Shetty,J., Shateman,S., Teegave,G., Geer,K.,
              Shvartsbeyn,A., Gebregorgis,E., Overton,L., Russell,D., Chen,D.,
              Riggs,F., de Jong,P. and Fraser,C.M.
TITLE        Rat BAC End sequences from Library CHORI-230 EcoRI segment
JOURNAL      Unpublished (1999)
COMMENT      Other GSSs: CH230-232F19.TV
              Contact: Shaying Zhao
              Department of Eukaryotic Genomics
              The Institute for Genomic Research
              9712 Medical Center Dr., Rockville, MD 20850, USA
              Tel: 301 838 0200
              Fax: 301 838 0208
              Email: szhao@tigr.org
              Clones are derived from the rat BAC library CHORI-230
              (http://www.chori.org/bacpac/rat230.htm). For BAC library
              availability, please contact Pieter de Jong (pdejong@mail.cho.org).
              Clones may be purchased from BACPAC Resources
              (http://www.chori.org/bacpac/or ering information.htm). BAC end
              page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
              Plate: 232 row: F column: 19
              Seq primer: SP6
              Class: BAC ends.
              Location/Qualifiers
                  1. 838
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /strain="BN/SENHsd/MCW"
                     /db_xref="taxon:10116"
                     /clone="CH230-232F19"
                     /sex="Female"
                     /cell_type="Brain"
                     /clone_lib="CHORI-230 Segment 1"
                     /note="vector: pPARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
                     CHORI-230 Rat (BN/SENHsd/MCW) BAC library produced by
                     Pieter de Jong"

FEATURES
             source
             1. 838
                Query Match      15.1%; Score 345.099; DB 1; Length 838;
                Best Local Similarity 67.1%; Pred. No. 20;
                Matches 557; Conservative 0; Mismatches 254; Indels 19; Gaps 5;

Qy 1213 CCTCTTTGATTTTGGCGCTGGAAATATTTATTATTCATATTTTCTTGAATGTGGGTAAAC 1272
Db 836 CATTCCTGTTGTTGCGCCCAAGACGATTAGTTCCTGCTTTTCTCTAGGGTATAGCTGCC 777
Qy 1273 ATCTTTAGATTGAAGTTTTTCTCTAGCCTCTT--TAGGTCTGCATTTGAAGATAGATA 1330
Db 776 CTCCTTATGCGGGCTTTACCCTTTATTATCCTTTGTAGTCTGGATTTGTAGAAAGATA 717
Qy 1331 TTCTTTACATCTGATTTTATCTTAGAATGCTCTTCTTCTCCAACTATTGTGACAGAAAG 1390
Db 716 TTGCTAAATTTGGTTTGTGTCATGGAATATCTTGTTTCCATCTATGTAATTGAGAG 657
Qy 1391 TTTTCTTAAGTGCAGTAGTCTGGCGCTGCACATCTCTAGTCTCTGGAGTCTGTAGCACATC 1450
Db 656 TTTTGCAGGATACAGTAACTCGGCGTGCATTTCTGTCTCTTAGGGTCTGTATGACATC 597
Qy 1451 TGTGAGGCGCTCTTACATTTTCAGTTTCTTATTTGGAAAGTCAAGGTGTAATTCATAATC 1510
Db 596 TGTCCAGGATCTTCTGGCGCTTCATAGTTTCT-GSCGAAAGTCTGTTGATCTGTATG 538
Qy 1511 ATCTGCCCTTATATGTTAAATGGTCTTTTTCCTTTGCACTTTTAATATCTTCTTTCTTG 1570
Db 537 GTCCTGCCCTTATATGTTACTT-GACCTTTTCCCTTACTGCTTTTAATATCTTCTTTCT 479
Qy 1571 TTCTATACCTTTTAGTGATTTTCATTTATATGCACTGTGGGAGTTTCTTTTCCGGTCCAA 1630
Db 478 TTTCTGTGTTGGTGTGTTTGACAAATATGACGGTAGGTGTTTCTTTCTGTCCTACT 419
Qy 1631 CTATTTGGTGTTTGTATGCTTCTGTACCTGTATGAGCACTCTCTTCTCAAGTTTAGGA 1690
Db 418 CTATTTGAGTCTCTAGGCTCTTGTATGCTATGGGTATCTCTTTTTTTAGGTAGGG 359

```

Qy	1691	AATTTTCCTTTTTGGTTCCTTGAAATAATTTCCCTGCTTTTGACCTG-----	1741
Db	358	AAGTTTCCTTAAGATTGTGTGAAGATAATTTACTGGCCTTTGAGCTGGGAGTCTTCA	239
Qy	1741	----CCCTTCCTCCCCTTCCTCTAATTCCTTTGGTTTTTGCATAGTGTCTCGGCTTCCTGG	1796
Db	298	CTCTCTCHATACTATTAATTCCTTAGTTTGATCTCTCATTTGAGTCTCGGATTTCTCTGT	239
Qy	1797	ATGTTTTATGCCGGGATTAATTTTAGACCTTAACAATTTCTTTTGACCRAAGGTATCCATTTCT	1856
Db	238	ATGTTTTGGACCAAGTACGTTTTTCCGCTTPACATTAATCTTTGACAGTTGAGTCAAUGATT	179
Qy	1857	TCTATCTGTCTTCACTGCCTCGAGATCTCTCTCTATCTCTCTGTATCTCTGTATCTGTGTCAGTGAGG	1916
Db	178	TCTATGAATCTTCTGCTCTCGAGATCTCTCTTCCATCTCTTGTATCTGTGTGGTGAAG	119
Qy	1917	CTTGCTCTCGAGTCTGTGTGGGTCTTAATTTTTTC-ATTTCCAGATTTCTCTTCAGTT	1975
Db	118	CTTGATCTACAGTCTCTGCTCTTTCTTTTGTGATTTTCTATGTCCAGGTTGTTCATG	59
Qy	1976	TGGTTTTGTGTTAATTAATCTATTTCCATTTTCAGTCTCGGAATGTTTT	2025
Db	58	TGTTCTTTCTTGATGCTCTATTTCCATTTTAATTCCTTCACTGTTT	9

RESULT 42
BH258120/c

LOCUS CH230-112118.TU CHORI-230 Segment 1 Rattus norvegicus genomic clone
DEFINITION BH258120, genomic survey sequence.
ACCESSION CH230-112118
VERSION BH258120
KEYWORDS BH258120.1 GI:17160443
SOURCE GSS.
ORGANISM Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 849)
AUTHORS Zhao S., Shetty J., Shatsman S., Teegaye G., Geer K.,
Shvartsbeyn A., Gebregorgis E., Overton L., Russell D., Chen D.,
Kiggs F., de Jong P. and Fraser C.M.
TITLE Rat BAC End Sequences from Library CHORI-230 EcoRI segment
JOURNAL Unpublished (1999)
COMMENT Other GSSs: CH230-112118.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pjejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or ering_information.htm). BAC end
page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
Plate: 112 row: I column: 18
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1 . 849
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SNHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-112118"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 1"
/note=Vector: pTABAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CHORI-230 Rat (BN/SNHsd/MCW) BAC library produced by

DEFINITION CH230-163N10.TJ CHORI-230 Segment 1 Rattus norvegicus genomic clone
CH230-163N10, genomic survey sequence.

ACCESSION BH357910

VERSION BH357910.1 GI:17288644

KEYWORDS GSS.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 762)

AUTHORS Zhao,S., Shetty,J., Shatsman,S., Teegaye,G., Geer,K.,

Shvartsbeyn,A., Gebregorgis,E., Overton,L., Russell,D., Chen,D.,

Riggs,F., de Jong,P. and Fraser,C.M.

Rat BAC End Sequences from Library CHORI-230 EcoRI segment

Unpublished (1999)

Other_GSSs: CH230-163N10.TV

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the rat BAC library CHORI-230

(<http://www.chori.org/bacpac/rat230.htm>). For BAC library

availability, please contact Pieter de Jong (pdjong@mail.choi.org).

Clones may be purchased from BACPAC Resources

(http://www.chori.org/bacpac/or_ering_information.htm). BAC end

page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html

Plate: 163 row: N column: 10

Seq primer: SP6

Class: BAC ends.

FEATURES

source

1..762

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/strain="BN/SSHsd/MCW"

/db_xref="taxon:10116"

/clone="CH230-163N10"

/sex="Female"

/cell_type="Brain"

/clone_lib="CHORI-230 Segment 1"

/note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;

CHORI-230 Rat (BN/SSHsd/MCW) BAC library produced by

Pieter de Jong"

Query Match 15.1%; Score 343.899; DB 1; Length 762;
Best Local Similarity 70.2%; Pred. No. 22;
Matches 526; Conservative 0; Mismatches 206; Indels 17; Gaps 5;

QY 1301 CTCTTTAGCTGCAATTCAGATAGATATCTTTACATCTGATTTATCTTAGAATGT 1360

DB 753 CCCTTTAGCGCTGGATTTGTAGAAAGATATTGTGTAATTTGGTTTGTACGGAATAT 694

QY 1361 CTCTTTCTCCCACTATTGTGACAGAAAGTTTCTTAAGTCAGTAGTCGCGCTGACA 1420

DB 693 CTTAGTTTCTCCATCAATGTTAATTGGGAGTTTGCAGGATACAGTAACCTGGGCTGCA 634

QY 1421 TCTGTAGTCTCTTGGAGTCTGTAGACATCTGTGACGGCCCTCTTACATTTTGAGTTTC 1480

DB 633 TTGTGTCTCTTAAGTCTGTATGACATCTGTCCAGGAACATCTGGCTTTCATAGTCTC 574

QY 1481 TATTGGAAGTCAGGTAAATCTTAATACATCTGCTTTATATGTAAATGCTCTTTT 1540

DB 573 T-GGAGAGAGTCTGGTGTGATCTCATAGGCTGCTTTATGTACTT-GACCTTTT 516

QY 1541 TCCTTTGCATCTTTTAATATCTTTCTTTGTCTATACCTTTTAGTGATTTGATTTATG 1600

DB 515 TCCTTTACTGCTTTTAATATCTTTCTTTTATTTGAGCAGCTTGATGTTTGACCATATG 456

QY 1601 CACTGTGGGGAGTTCTTTTCCGGTCCAACTATTGTGGTGTGTTGTATGCTTCTGTACC 1660

DB 455 TGATGGGGGAGTTTCTTTTCTGGTCCAACTATTGGAGTCTGTAGGCTTCTTGTATG 396

QY 1661 TTGATAGGCATCTCTTTCTCAAGGTAGGAAATTTTCTTTTGGTTTCTTGAATA 1720

DB 395 TTTATGGGCATCTCTTTCTTTAGGTAGGGAAGTTTCTTCTATGATTTTGTGAAGATA 336

QY 1721 TTTTCCCTGCTTTTGACCTGC-CTTCTCCCTTCTCTATTC- 1766

DB 335 TTTACTGCTCTTTTGAGCTGGCGCTTCTCACTCTCTCTATACCTATATCCITAGGTTT 276

QY 1767 GGTTTTGCATAGTGTCTCTGGCTTCCCTGGATGTTTTATGCTCGATATTATTAAGACTTA 1826

DB 275 GATCTTCTCATAGTCTCTGGATTTCTGTATGTTTTGGACCACTAGCTTTTCTGTTT 216

QY 1827 ACATTTTCTTTTGACCAAGGTATCCATTTCTTCTATCTTGTCTTCTCACCTGCTGAGATTCTC 1886

DB 215 ACATTAICTTTTGACAGTTGTGTGATGATTTCTAAGGAATCTTCTCTCCGAGATTCTC 156

QY 1887 TCTTCTATCTCTTGTATTCTGTCTAGTGAGGCTTGTCTCTGAGTTCC-TGTGGGTTCTT 1945

DB 155 TCTTCTATCTCTTGTATTCTGTGTAATGCTTGTATCTATGCTCTCTGCTCTTCTCT 96

QY 1946 AATTTTTCATTTCCAGATTTCTTCAAGTTTGGGTTTGTGTTTATTAATTTCTATTTCCACT 2005

DB 95 TGGTTTTCTATATCCAGGTTGTCTTCCCTTTGCTTTTCTTTATTGCTTCTATTTCCATT 36

QY 2006 TTCAGGTCCTGAAATGTTTTACTCATTTT 2034

DB 35 TTTAATTCCTTCACCTGTTGCTTGCTT 7

Search completed: August 25, 2004, 09:33:44
Job time : 181 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 08:45:30 ; Search time 14588 Seconds
(without alignments)
3.946 Million cell updates/sec

Title: US-10-664-775-2

Perfect score: 3572

Sequence: 1 gtcaggagggcgcagtgta.....gcacacacgcagaagctt 3572

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 0.5

Searched: 44 seqs, 8057215 residues

Total number of hits satisfying chosen parameters: 88

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : rge2.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	2197.3	61.5	160990	1	AC141413
C 2	2158.67	60.4	233345	1	AC141413
C 3	2153.4	60.3	198916	1	AC145008
C 4	2147.7	60.1	179015	1	AC135839
C 5	2147.7	60.1	241923	1	AC137656
C 6	2131	59.7	173068	1	AC091660
C 7	2125.5	59.5	171153	1	AC135215
C 8	2123.49	59.4	145088	1	AC142564
C 9	2123.49	59.4	192431	1	AC141853
C 10	2108.1	59.0	173207	1	AC108887
C 11	2107.7	59.0	213549	1	AC140963
C 12	2106.5	59.0	196994	1	AC144887
C 13	2105.2	58.9	192199	1	AC139312
C 14	2103.2	58.9	145088	1	AC142564
C 15	2100	58.8	189278	1	AC142562
C 16	2095.4	58.7	224721	1	AC139632
C 17	2081.09	58.3	193383	1	AC091252
C 18	2071.6	58.0	139617	1	AC146889
C 19	2071.2	58.0	172022	1	AC145440
C 20	2066.1	57.8	294688	1	AC144498
C 21	2065.3	57.8	153264	1	AC135216
C 22	2065	57.8	223761	1	AC147195
C 23	2062.1	57.7	180432	1	AC134959
C 24	2052.6	57.5	195590	1	AC146537
C 25	2048.6	57.4	133710	1	AC093750
C 26	2046.4	57.3	181921	1	AC091399
C 27	2032.6	56.6	175949	1	AC107065
C 28	2018.9	56.5	177535	1	AC109914
C 29	1982.6	55.5	207875	1	AC147592
C 30	1945.1	54.5	192540	1	AC137651
C 31	1931.7	54.1	163604	1	AC092727
C 32	1930.2	54.0	168497	1	AC092410
C 33	1919.1	53.7	173952	1	AC130787

ALIGNMENTS

RESULT 1

AC141413/c

LOCUS

DEFINITION

AC141413

Accession

Version

Keywords

Source

Organism

Reference

Authors

Title

Journal

Reference

Authors

Title

Journal

Comment

AC141413 160990 bp DNA linear HTG 07-MAY-2003
Bos taurus clone RP42-99E14, WORKING DRAFT SEQUENCE, 8 ordered pieces.

AC141413

AC141413.2 GI:30409902

HTG: HTGS_PHASE2; HTGS_DRAFT.

Bos taurus (cow)

Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovidae; Bovinae; Bos.

1 (bases 1 to 160990)

Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,

Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Cariaga, K.,

Chu, G., Coleman, B., Coleman, H., Engle, J., Granite, S., Guan, X.,

Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Hu, P.,

Hurle, B., Idol, J.R., Karlins, E., Kwong, P., Latic, P., Lee-Lin, S.-Q.,

Legaspi, R., Maduro, Q.L., Maduro, V.B., Margulies, E.H., Masello, C.,

Maskeri, B., McDowell, J., Paguirigan, C., Pearson, R., Portnoy, M.E.,

Prasad, A., Reddix-Dugue, N., Schandler, K., Schueler, M.G., Shah, K.,

Sison, C., Stantripp, S., Thomas, J.W., Thomas, P.J., Tsipouri, V.,

Vogt, J.L., Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.

NISC Comparative Sequencing Initiative

Unpublished

2 (bases 1 to 160990)

Green, E.D.

Direct Submission

Submitted (14-MAR-2003) NIH Intramural Sequencing Center, 8717

Grovesmont Circle, Gaithersburg, MD 20877, USA

3 (bases 1 to 160990)

Green, E.D.

Direct Submission

Submitted (07-MAY-2003) NIH Intramural Sequencing Center, 8717

Grovesmont Circle, Gaithersburg, MD 20877, USA

On May 7, 2003 this sequence version replaced gi:28951185.

----- Genome Center

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: <http://www.nisc.nih.gov>

Contact: nisc.zoonhgri.nih.gov

----- Project Information

Center project name: edv

Center Clone name: 099E14

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence

contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics -----
 Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 159861 bases at least Q40
 Consensus quality: 160028 bases at least Q30
 Consensus quality: 160216 bases at least Q20
 Insert size: 162000; agarose-fp
 Insert size: 160290; sum-of-contigs
 Quality coverage: 11.34x in Q20 bases; agarose-fp
 Quality coverage: 11.46x in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

1 50852: contig of 50852 bp in length
 50853 50952: gap of unknown length
 50953 87597: contig of 36645 bp in length
 87598 87697: gap of unknown length
 87698 90744: contig of 3047 bp in length
 90745 90844: gap of unknown length
 90845 97904: contig of 7060 bp in length
 97905 98004: gap of unknown length
 98005 106019: contig of 8015 bp in length
 106020 106119: gap of unknown length
 106120 130293: contig of 24174 bp in length
 130294 130393: gap of unknown length
 130394 156781: contig of 26388 bp in length
 156782 156882: gap of unknown length
 156882 160990: contig of 4109 bp in length.

FEATURES

source

1. 160990
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9913"
 /clone="RP42-99814"
 /clone_lib="RP42"

misc_feature

1. 50852
 /note="assembly_fragment
 clone end:SP6"

misc_feature

vector_side:left

misc_feature

50953..87597
 /note="assembly_fragment"

misc_feature

87698..90744
 /note="assembly_fragment"

misc_feature

90845..97904
 /note="assembly_fragment"

misc_feature

98005..106019
 /note="assembly_fragment"

misc_feature

106120..130293
 /note="assembly_fragment"

misc_feature

130394..156781
 /note="assembly_fragment"

misc_feature

156882..160990
 /note="assembly_fragment
 missing T7 clone end on 3' end of insert"

Query Match

Best Local Similarity 61.5%; Score 2197.3; DB 1; Length 160990;
 Matches 2635; Conservative 77.5%; Pred. No. 0;

1 GTCAGGAGGGCGGCGAGTGGAGGAGGTACTACCTCGTCCAGGTAAAGGAGCAGTAGC 60

QY

115101 GTCAGGAGGGTGGCGGTGAGGAGA-----TACCCTCGTCCAGGTAAAGTACGACGC 115047

Db

QY	61	TCGGCTTGTCTGGAGCAGCGGTAAAGAGATACCCACGCCCCAAGGTAAAGAAACCCCAAG	120
Db	115046	TCGGCTTGTCTGGAGCAGCGGTAAAGAGATACCCACGCCCCAAGGTAAAGAAACCCCAAG	114987
QY	121	TAAGATGTGTAGGTGTGTGAGAGGCGCATCAGAGGCGCAGACATCTCTGAAACCCATACAGCA	180
Db	114986	TAAGACGGTGTAGGTGTGTGAGAGGCGCATCAGAGGCGCAGACATCTCTGAAACCCATACAGCA	114927
QY	181	GAATACTAGTCTAATCTAATCACTAGACACACAGCCTTGTCTAATCAATGAACCTAAG	240
Db	114926	GAATACTAGTCTAATCTAATCACTAGACACACAGCCTTGTCTAATCAATGAACCTAAG	114867
QY	241	CCATGCCGTGGGGCAACCCAAAGATGGGCGGTCTATGTTGGAGAGATCTGACAAATGTG	300
Db	114866	CCATGCCGTGGGGCAACCCAAAGATGGGCGGTCTATGTTGGAGAGATCTGACAAATGTG	114807
QY	301	GTCCACTGGAGAGGGAATGCAAACTTCTAGTATTCTTCCCTTGAACCCCATGAAC	360
Db	114806	GTCCACTGGAGAGGGAATGCAAACTTCTAGTATTCTTCCCTTGAACCCCATGAAC	114747
QY	361	AGTATGAAGGCAAAATGATAGGATCTGAAAGAGGAACCTCCAGGTCTAGTAGGTGCC	420
Db	114746	AGTATGAAGGCAAAATGATAGGATCTGAAAGAGGAACCTCCAGGTCTAGTAGGTGCC	114687
QY	421	CCATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGATGAAGAGATGGAGCA	480
Db	114686	CAATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGATGAAGAGATGGAGCA	114627
QY	481	RAGCAAAAGATATACCCAGCTGTGGATGTGATGTTGATGATAGAGGTCCGATGCTGT	540
Db	114626	RAGCAAAAGATATACCCAGCTGTGGATGTGATGTTGATGATAGAGGTCCGATGCTGT	114567
QY	541	AAAGAGCAATATTGTCATAGGAACCTGGAATCTCAGGTCCATGAATCAAGGCAAAATGGAA	600
Db	114566	AAAGAGCAATATTGTCATAGGAACCTGGAATCTCAGGTCCATGAATCAAGGCAAAATGGAA	114507
QY	601	GTGGTCAAAAGAGATGGCAAGAGTGAATGTCAACATCTTAGGAATCAGCGAACTAAAA	660
Db	114506	GTGGTCAAAAGAGATGGCAAGAGTGAATGTCAACATCTTAGGAATCAGCGAACTAAAA	114447
QY	661	TGGACTGGAATGGGTGAATTTAACTCAGATGACCATATATCTACTCTCGGCGGAGAA	720
Db	114446	TGGACTGGAATGGGTGAATTTAACTCAGATGACCATATATCTACTCTCGGCGGAGAA	114387
QY	721	TCCCTCAGAGAAATGGAGTAGCCATCATGTCACAAAAGAGTCCGAAATGCAAGTACTT	780
Db	114386	TCCCTCAGAGAAATGGAGTAGCCATCATGTCACAAAAGAGTCCGAAATGCAAGTACTT	114327
QY	781	GGATCGACTCAAAAACGACAGATGATCTCTGTTTGTTCAGGCAACCAATTCAT	840
Db	114326	GGATCGACTCAAAAACGACAGATGATCTCTGTTTGTTCAGGCAACCAATTCAT	114267
QY	841	ATCAGATTAATCCAAAGTCTATGCCCCCAACCAAGTAACTGCTGAAGAGGTGAAGTCGG	900
Db	114266	ATCAGATTAATCCAAAGTCTATGCCCCCAACCAAGTAACTGCTGAAGAGGTGAAGTCGG	114207
QY	901	TCCCTAGAGACCTTACAGACCTTTTGAAGTAACTACCCCAAAAAGATGCTCTCTCAT	960
Db	114206	TCCCTAGAGACCTTACAGACCTTTTGAAGTAACTACCCCAAAAAGATGCTCTCTCAT	114147
QY	961	ATAGGGGACTGGAATGCAAAAGTAGGAAGCAAAAACACCTGGAGTAAACAGGCAAAATTT	1020
Db	114146	ATAGGGGACTGGAATGCAAAAGTAGGAAGTAAAGAAACCTGGAGTAAACAGGCAAAATTT	114087
QY	1021	GGCTTGAATACCGAATGAAGAGGCGCAAGACTAATAGAGTTTCCCAAGAAATGCA	1080
Db	114086	GGCTTGAATACCGAATGAAGAGGCGCAAGACTAATAGAGTTTCCCAAGAAATGCA	114027
QY	1081	CTGTGTATAGCAAAACACCTCTTCCCAACACACAGAGAGAGCTCTACACATGACATCA	1140
Db	114026	CTGTGTATAGCAAAACACCTCTTCCCAACACACAGAGAGAGCTCTACACATGACATCA	113967
QY	1141	CCAGATGGTCAACACCGGAAATTCAGATTGATTATTTCTTTCGACGCCAAGATGGGAAGC	1200

Db 113966 CCAGATGTCACACCGAAATCAGATTGATTATATCTTCGACGCAAGATGAGAGC 113907
Qy 1201 TCTATACAGTCAGCAAAACAGACACAGAGCTTACTGTGGCTCAGATCATGAATCCTT 1260
Db 113906 TCTATACAGTCAGCAAAACAGACACAGAGCTGACTGTGGCTCAGATCATGAATCCTT 113847
Qy 1261 ATTGCCAAATTCAGACTTAAATTGAAGAAAGTAGGGAACCACTAGATCATCAGGTAA 1320
Db 113846 ATAGCCAAATTCAGACTTAAATTGAAGAAAGTAGGGAACCACTAGATCATCAGGTAT 113787
Qy 1321 GACCTAAATCCAAATCCCTTATGATATACAGTGAAGTAGAGAAATAGATTTAAGGGCCTA 1380
Db 113786 GACCTAAATAAATTCCTTATGATCATACAGTGAAGTAGAGAAATAGATTTAAGGGCCTA 113727
Qy 1381 GATCTGATAGACAGAGTACTTAATGAACATATGACAGAGTTTCATGATGTACAGAG 1440
Db 113726 GATCTGATAGATAGAGTACTTGAATGAACTATGGAATGAGTTTCGATGATGTATATGAG 113667
Qy 1441 ACAGGATCGACACCATCCCAATGGAAGAAAGAAATGCAAAAGCAAAATGGCTGTCTGG 1500
Db 113666 ACAGGATCAAGACCATCCCAATGGAAGAAAGAAATGCAAAAGCAAAATGGCTGTCTGG 113607
Qy 1501 GAGGCTTACAAATAGCTGTGAAAAGAGAGAAAGTGAAGCAAAAGGAAAGAAAGA 1560
Db 113606 GAGGCTTACAAATAGCTGTGAAAAGAGAGAAAGTGAAGCAAAAGGAAAGAAAGA 113547
Qy 1561 TAAAGCATCTGAATCAGAGTTCCAA 1588
Db 113546 TATAAGCATCTGAATCAGATTTCAGAGAAATAGCAAGAGAGATGAAGAAAGCTTCCTC 113487
Qy 1589 1588
Db 113486 AGCGATCAATGCAAGAAATAGAGGAAACACAGNAATGGGAAGACTAGAGATCTCTTC 113427
Qy 1589 1588
Db 113426 AAGAAAATAGAGATACCAAGGAACATTTTCAGCAAGATGGGCTCGATAAAGACAGA 113367
Qy 1589 1588
Db 113366 AATGTTAGGACCTAACAGAGCAGAGAGATATGAAGAGAGATGGCAAGATACACAGAA 113307
Qy 1589 1588
Db 113306 GAACTGTACAAAAGATCTTACAGAGCCAGATAATCACGATGGTGTGATCACTGACCTA 113247
Qy 1589 1588
Db 113246 GAGCCAGACATCTTGGAAATGTGAAGTCAAGTGGGCTTGAAGAGCATCTACAAACAAA 113187
Qy 1589 1588
Db 113186 GCTAGTGGAGTGTGAATTCAGGTGAGCTATTTCAATCTGAAAGATGATGCTGTG 113127
Qy 1589 1588
Db 113126 AAGTGTGCTGACCTCAATATGCCAGAAATCTGGAAGAACTCAGCAGTGGCCACAGGACTGG 113067
Qy 1589 1588
Db 113066 AAAAGTCAATTTTCATTCCAAATCCCAAGAAAGGCAATGCCAAAGAAATGCTCAAACTAC 113007
Qy 1589 1588
Db 113006 CGCACAGTTCGACTCATCTCACAGCTAGTAAAGTAATGCTCAAAATTTCTCCAAGCCAGG 112947
Qy 1589 1588
Db 112946 CTTCAGCAAGATTTGAAGTGTGAATCTTCTGATGTTCAAGCTGTGTTTGAAGAGGAGG 112887
Qy 1628 GGAACAGAGACCAATTTGCCAATCTCTGTATCATGGAAGAAAGCAAGAGAGTTCCAG 1687

Db 112886 GGAACAGAGATCAAAATTCGCTCCAACTTTGCTGGATCATAGAAAAAGCAAGAGAGTTCCAG 112827
Qy 1688 AAAAAATCTATTTCTGCTTTTATGATATGCAAAAGCCTTTTCACTGTGGGCTCAAT 1747
Db 112826 AAAAAATCTATTTCTGCTTTTATGATATGCAAAAGCCTTTTCACTGTGGGCTCAAT 112767
Qy 1748 AAATCTGTGGAATTTCTGAAAGGATGGGAATACCAAGACCACTGACCTGACTCTTTGAAA 1807
Db 112766 AAATCTGTGGAATTTCTGAAAGGATGGGAATACCAAGACCACTGACTCTGCTTTGAGA 112707
Qy 1808 AATTTGATCAGGTCAGGAAGCAACGTTAGAACTGGACATGGAACACAGACTGGTTTC 1867
Db 112706 AATCTATGAGTCAGGAAGCAACGTTAGAACTGGACATGGAACACAGACTGGTTTC 112647
Qy 1868 CAAGTGAAGAAAGGATGATGTCAGGCTGTATATGTCACCCGCTGTGTTTAACTTTCTAT 1927
Db 112646 CAATAGGAAGAGGATGTCAGGCTGTATATGTCACCCGCTGTGTTTAACTTTATAT 112587
Qy 1928 GCAGAG-ACATCATAGAAAACGCTGGGCTGGAAAGAACACAAGCTGGAATCAAGATTGCC 1986
Db 112586 GCAGATACATCATGAGAAAACCTGGACTGGAAGAAACACAAGTTGGAATCAAGATTGCC 112527
Qy 1987 GGGAAATAGCAATTAACCTCAGATATGCAGATGATACCACTTATGCGCAGAAAAGTGA 2046
Db 112526 GGGAAATATCAATAACCTCAGATATGCAGATGATACCACTTATGCGCAGAAAAGTGA 112467
Qy 2047 GAGAACTTAAAGCCTCTTTCATGAAGTGAAGAGGAGAGTGAAGAAAGTTGGCTTAAAG 2106
Db 112466 GAGAACTTAAAGCCTCTTTCATGAAGTGAAGAGTGAAGAAAGTTGGCTTAAAG 112407
Qy 2107 CTCAACTTGAAGAAACAGAGATCATGSCATCTGGTCCCATCACTTTCATGGGAATAGAT 2166
Db 112406 CTCAACTTGAAGAAACAGAGATCATGSCATCTGGTCCCATCACTTTCATGGGAATAGAT 112347
Qy 2167 GGGAAACAGTGGAAACAGTGTGAGCTTTATTTTGGGGGGCTCCAAAATCACTGCAG 2226
Db 112346 GGGAAAGATGGAAACAGTGTGAGCTTTATTTTGGGGGGCTCCAAAATCACTGCAG 112287
Qy 2227 ATGTTGATCGAGCATGAAATTAAGACACTTACTCTCTTGGGAAGAAAGTTA-ACCA 2284
Db 112286 ATGTTGATCGAGCATGAAATTAAGACACTTACTCTCTTGGGAAGAAAGTTAAGACCA 112227
Qy 2285 ACCTAGATAGCATATGGAAGCAGAGACATTTACTCTGCCAAAGAAAGCCCATCTAGTCA 2344
Db 112226 ACCTAGATAGCATATGGAAGCAGAGACATTTACTCTGCCAAAGAAAGTTGACTAGTCA 112167
Qy 2345 AGGCTATGCTTTTCCAGTGTGATGATGATGAGTGTGAGTGTGAGTGTGAAGAAAGCTG 2404
Db 112166 AGGCTATGCTTTTCCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 112107
Qy 2405 AGCATGAAGAAATGATGCTTTTGAACGTGCTGTGAGAGAGACTCTTGAAGTGTGCTT 2464
Db 112106 AGCGCAAGAAATGATGCTTTTGAACGTGCTGTGAGAGAGACTCTTGAAGTGTGCTT 112047
Qy 2465 GGACTGCAAGGAGATCCCAACAGTCCATTTCTGAAGGAGATCAGCCCTGGGATTTCTTTGG 2524
Db 112046 GGACTGCAAGGAGATCCCAACAGTCCATTTCTGAAGGAGATCAGCCCTGGGATTTCTTTGG 111987
Qy 2525 AAGGAATGATGCTAAAGCTGAAAACCTCCAGTACTTTTGGCCACTCATGGAAGAGTGTGACT 2584
Db 111986 AAGGAATGATGCTAAAGCTGAAAACCTCCAGTACTTTTGGCCACTCATGGAAGAGTGTGACT 111927
Qy 2585 CACTGGAAGAAAGCCTGATGCTGGAGGATTTGGGGCAGGAGAGAGAGGAGGAGCAGAGA 2644
Db 111926 CATTGGAAGAAAGCTCTGATGCTGGAGGATTTGGGGCAGGAGAGAGAGGAGGAGCAGAGA 111867
Qy 2645 GGATGAGATGGCTGATGGCATCACTGATCGATGAGCTGAGTGTGGTGAATCTCTGG 2704
Db 111866 GGATGAGATGGCTGATGGCATCACTGATCAATGAGCTGAGTGTGGTGAATCTCTGG 111807
Qy 2705 AGTTGATGATGAGAGGAGGCTGCTGCTGGCCATTCATGGGCTCAAGAGAGTGTGGA 2764
Db 111806 AGTTGATGATGAGAGGAGGCTGCTGCTGGCCATTCATGGGCTCAAGAGAGTGTGGA 111747

```

QY      2765  CACGACTGAGCAACTGAACCTGAACCTGAACCTGACTGAA 2802
      |||||||
Db      111746  CATGACTGAGCAACTGAACCTGAACCTGAACCTGATGGGA 111709

RESULT 2
AJ421481/c      AJ421481      233345 bp      DNA      linear      MAM 12-JUN-2002
LOCUS      Bos taurus X-inactivation center region, Jpx and Xist genes.
DEFINITION
ACCESSION      AJ421481
VERSION
KEYWORDS      Jpx gene; Xist gene.
SOURCE      Bos taurus (cow)
ORGANISM      Bos taurus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.

REFERENCE
1 Chureau,C., Prissette,M., Bourdet,A., Barbe,V., Cattolico,L.,
  Jones,L., Eggen,A., Avner,P. and Duret,L.
  Comparative sequence analysis of the X-inactivation center region
  in mouse, human, and bovine
  Genome Res. 12 (6), 894-908 (2002)
JOURNAL
MEDLINE      22040331
PUBMED      12045143
REFERENCE
2 (bases 1 to 233345)
Duret,L.
Direct Submission
Submitted (27-NOV-2001) Duret L., Biologie et Biologie Evolutive,
CNRS - Universite Lyon 1, 16, rue Raphael Dubois, 69622
Villeurbanne cedex, FRANCE
COMMENT
Sequenced by the Genome. Contact: Philip Avner, Institut
Pasteur, 25, rue du Docteur Roux, 75015 Paris, France.
FEATURES
source
1. .233345
   /organism="Bos taurus"
   /mol_type="genomic DNA"
   /db_xref="taxon:9913"
   complement(49..107)
   /rpt_family="MLT1B LTR/MaLR"
   complement(129..519)
   /rpt_family="MLT1C LTR/MaLR"
   complement(527..1161)
   /rpt_family="L1M4 LINE/L1"
   complement(1357..1531)
   /rpt_family="Bov-tai SINE/BovA"
   complement(1594..1868)
   /rpt_family="L1M4 LINE/L1"
   2792..2903
   /rpt_family="L1M4a LINE/L1"
   complement(2965..3169)
   /rpt_family="Bov-B LINE/BovB"
   3193..3423
   /rpt_family="L1M4a LINE/L1"
   complement(3413..3539)
   /rpt_family="L1Mdb LINE/L1"
   3548..3642
   /rpt_family="MERS8 DNA/MER1_type"
   3601..3984
   /notes="At rich low_complexity"
   4259..4293
   /notes="At rich low_complexity"
   complement(4375..5337)
   /rpt_family="L1ME LINE/L1"
   5360..5799
   /rpt_family="L1M4a LINE/L1"
   5810..5912
   /rpt_family="CHR-2 SINE/whales"
   6143..7596
   /rpt_family="Bov-B LINE/BovB"
   7831..7941
   /rpt_family="L1M4a LINE/L1"

repeat_region      complement(7957..8179)
repeat_region      /rpt_family="Bov-tai SINE/BovA"
repeat_region      8323..8443
repeat_region      /rpt_family="MERS8 DNA/MER1_type"
repeat_region      8445..8713
repeat_region      /rpt_family="Bov-A2 SINE/BovA"
repeat_region      8974..9118
repeat_region      /rpt_family="L1M4a LINE/L1"
repeat_region      9164..9391
repeat_region      /rpt_family="L1M4a LINE/L1"
repeat_region      9448..9486
repeat_region      /note="GA-rich Low_complexity"
repeat_region      9624..10180
repeat_region      /rpt_family="L1M4a LINE/L1"
repeat_region      10484..10571
repeat_region      /rpt_family="L1MD3 LINE/L1"
repeat_region      10557..10676
repeat_region      /rpt_family="Bov-ta2 SINE/BovA"
repeat_region      10766..11017
repeat_region      /rpt_family="L1M4 LINE/L1"
repeat_region      11069..11339
repeat_region      /rpt_family="L1_B LINE/L1"
repeat_region      11520..11680
repeat_region      /rpt_family="L1ME LINE/L1"
repeat_region      12070..12235
repeat_region      /rpt_family="L1_Art LINE/L1"
repeat_region      13218..13257
repeat_region      /rpt_family="L1M4a LINE/L1"
repeat_region      13258..13282
repeat_region      /note="(CA)n Simple_repeat"
repeat_region      13283..13318
repeat_region      /rpt_family="L1M4a LINE/L1"
repeat_region      complement(13342..13717)
repeat_region      /rpt_family="Bov-B LINE/BovB"
repeat_region      14072..14450
repeat_region      /rpt_family="L1M4a LINE/L1"
repeat_region      15056..15154
repeat_region      /rpt_family="MER34 LTR/ERV1"
repeat_region      complement(15227..15343)
repeat_region      /rpt_family="HAL1b LINE/L1"
repeat_region      15648..15697
repeat_region      /rpt_family="ARLTR2 LTR"
repeat_region      15824..15954
repeat_region      /rpt_family="ARLTR2 LTR"
repeat_region      complement(16049..16155)
repeat_region      /rpt_family="HAL1b LINE/L1"
repeat_region      complement(16585..16726)
repeat_region      /rpt_family="HAL1b LINE/L1"
repeat_region      complement(16834..16994)
repeat_region      /rpt_family="Bov-ta2 SINE/BovA"
repeat_region      17795..18827
repeat_region      /rpt_family="L1MB8 LINE/L1"
repeat_region      18845..19964
repeat_region      /rpt_family="L1MB8 LINE/L1"
repeat_region      complement(20011..20113)
repeat_region      /rpt_family="Bov-D SINE/BovD"
repeat_region      20114..20192
repeat_region      /rpt_family="L1MB8 LINE/L1"
repeat_region      20190..20311
repeat_region      /rpt_family="L1MB8 LINE/L1"
repeat_region      20353..20898
repeat_region      /rpt_family="L1MB8 LINE/L1"
repeat_region      20903..21264
repeat_region      /rpt_family="Bov-A2 SINE/BovA"
repeat_region      21345..21593
repeat_region      /rpt_family="L1MB8 LINE/L1"
repeat_region      21598..21898
repeat_region      /rpt_family="L1Mda LINE/L1"
repeat_region      21899..21936
repeat_region      /note="(TA)n Simple_repeat"
repeat_region      21937..22365
repeat_region      /rpt_family="L1Mda LINE/L1"
repeat_region      22486..22767

```


QY 1198 AGCTCTATACAGTACGCAAAACCAAGACCAGAGCTTACTGTGGCTCAGATCATGAATC 1257
Db 222780 AGTTCTATACAGTACGCAAAACCAAGACCAGAGCTGACTGTGGCTCAGATCATGAATC 222781
QY 1258 CTTATTCGCAATTCAGACTTAAATTCGAAGAAGTAGGAAACCACTAGATCACTCAGG 1317
Db 222720 CTTATTCGCAATTCAGACTTAAATTCGAAGAAGTAGGAAACCACTAGATCACTCAGG 222661
QY 1318 TAAGACCTAAATCCATCCCTTTATGATTATATACAGTGGAAAGTAGAAATAGATTAAAGGGC 1377
Db 222660 TATGATCTAAATCAATCCCTTTATGATTATAGAGT-----GATTAAAGGGC 222615
QY 1378 CTAGATCTGATAGACAGAGTACCTAATGAACCTATGACAGAGTTGATGACATGTTACAG 1437
Db 222614 CTAGATCTGATAGATAGAGTGCCTGATGAATGAACTATGGAATGAGTTGATGACATGTTACAG 222555
QY 1438 GAGACAGGGATCGAGACCATCCCATGGAAAGAAATGCAAAAAAGCAAAATGGCTGTCT 1497
Db 222554 GAGACAGGGATCAAGACCATCCCATGGAAAGAAATGCAAAAAAGCAAAATGGCTGTCT 222495
QY 1498 GGGAGGCCCTTACAAATAGCTGTGAAAGAGAGAGTGAAGAACAAAGCAAAAGGAA 1557
Db 222494 GGGAGGCCCTTACAAATAGCTGTGAAAGAG--AGAAAGCGAAAGCAAAAGGAAAGGAA 222437
QY 1558 AGATAAAGCATCTGATGACAGAGTCCAAAG----- 1589
Db 222436 AGATAAAGCATCTGATGACAGAGTCCAAAGAGATGCAAGAGAGATAAGAAAGCCTTC 222377
QY 1590 ----- 1589
Db 222376 TTCAGTGATCAATGCAAGAAATAGAGGAAACAAACAGATGGAAGACTAGAGATCTC 222317
QY 1590 ----- 1589
Db 222316 TTCAGAAATATAGATACCAAGGAAACATTTCTGCAAAAGATGGGCTTGATAAGGAC 222257
QY 1590 ----- 1589
Db 222256 AGAATGGTATGGCCCTACAGAAAGCAGAGATATTAAAGAGAGTGGCAAGATACACC 222197
QY 1590 ----- 1589
Db 222196 AAAGAACTGTCAAAAAAGATCTTACAGCCAGATAATCAAAATGTCACTGAGCTAGAG 222137
QY 1590 ----- 1589
Db 222136 CCAGACATCTGGAATGTGAAGTCAAGTGGGCTTAGAAATCATCTACGAACAAAGCT 222077
QY 1590 ----- 1589
Db 222076 AGTAGAGTGATGGAATTCAGTTGAGTACTTCAATCTGAAAGATGATGCTGTGAAA 222017
QY 1590 ----- 1589
Db 222016 GTGCTGCATCAATATGCCAGCAATTTGGAAAACTCAGCAGTGGCCAGAGACTGGAAA 221957
QY 1590 ----- 1589
Db 221956 AGGTGAGTTTTCATTCCAATCCCAAGAAAGGAATGCCAAACAATGCTCAAACTACCAC 221897
QY 1590 ----- 1589
Db 221896 ACAATGCACTCATCAACGCTAGTAAAGTAATGCTCAATGTCTCCAAAGCCAGGCTT 221837
QY 1590 -----AACTTCAGTTGTTCAGCTGGTTTAGAAAAAGTACAGAGA 1630
Db 221836 CAGCAATATGTGAACGTTAACTTCCAAATGTTCAAGCTGTTTTAGAAAGGCAAGAGA 221777
QY 1631 ACCAGAGCAAAATTCGCAACATCTCTGTATATCTTGGAAAAAGCAAGAGAGTTCCAGAAA 1690
Db 221776 ACCAGAGATCAAAATTCGCAACATCCGCCAGATCATCAAAAAAGCAAGAGAGTTCCAGAAA 221717

QY 1691 AACATCTATTTCTGCTTTATTTGACTATGCAAAAGCCTTTGACTGTGGGGTCACAATAAA 1750
Db 221716 AACATCTATTTCTGCTTTATTTGACTATGCAAAAGCCTTTGACTGTGGATCACAATAAA 221657
QY 1751 CTGTGGAAAAATTCGAAAGGGATGGGAATACAGACCACTGACTGACTCTTTGAAAAAT 1810
Db 221656 CTGTGGAAAAATTCGAAAGAGATGGGAATACAGACCACTGACTGACTCTTTGAAAAAT 221597
QY 1811 TTGTATGACAGTTCAGAAAGCAACAGTTAGAACTGGACATGAAACAAACAGACTGTTCCAA 1870
Db 221596 TTGTATGACAGTTCAGAAAGCAACAGTTAGAACTGGACATGAAACAAACAGACTGTTCCAA 221537
QY 1871 GTAGAAAAAGGAGTATGTCAAGGCTGTATATTTGACCCCGCTTTTAACTCTTATGCA 1930
Db 221536 ATAGAAAAAGGAGTATCATCAAGGCTGTATATTTGACCCCTGTTTAACTCTTATGCA 221477
QY 1931 GAG-ACATCATGAGAAACGCTGGCTGGAAGCAAGCAAGCTGGAATCAAGATTGCCGG 1989
Db 221476 GAGTATATCATGAGAAACGCTGGCTGGAAGCAAGCTGGAATCAAGATTGCCAGG 221417
QY 1990 AGAAATACAAATAACCTCAGATATTCAGATGATACCACTTATGCGAGAAAGTGAAGAG 2049
Db 221416 AGAAATATCAATAACCTCAGATATTCAGATGATACCACTTATGCGAGAAAGTGAAGAG 221357
QY 2050 GAACTTAAAGCCCTTTGATGAAGGTGAAGAGAGAGTGAAGAAAGTGGCTTAAAGCTC 2109
Db 221356 GAACTTAAAGCCCTTTGATGAAGGTGAAGAGAGAGTGAAGAAAGTGGCTTAAAGCTC 221297
QY 2110 AACATTCAGAAAAAGAGATCATGGCATCTGGTCCCATCACTTATGCGAAATAGATGGG 2169
Db 221296 AACATTCAGAAAAAGAGATCATGGCATCTGGTCCCATCACTTATGCGAAATAGATGGG 221237
QY 2170 GAAACAGTGGAAACAGTGTGAGCTTTATTTTGGGGGGCTTCAAAATCACTGCGAGATG 2229
Db 221236 GAAACAGTGGAAACAGTGTGAGATTTA-TTTTGGGGGGGTTCAAAATCACTGCGAGATG 221178
QY 2230 GTGACTGAGCCATGAATTTAAAGAGACACTTACTCTTGAAGAAAGTTA--ACCAACC 2287
Db 221177 GGGACTGAGCCATGAATTTAAGAGAGCTTACTCTTGAAGAAAGTTAAGCAACCC 221118
QY 2288 TAGATAGCATATTGAAAGCAGAGACATTAACCTTGGCAACAAAGCCCACTAGTCAAGG 2347
Db 221117 TAGATAGCATATTCAAAAGCAGAGACATTAACCTTGGCAACAAAGCTGTCTAGTCAAGG 221058
QY 2348 CTATGGTTTTTCCAGTGGTCTGATGATGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2407
Db 221057 CTATGGTTTTTCCAGTGGTCTGATGATGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 220998
QY 2408 ACTGAAGAAATGATGCTTTTGAACCTGTGGTGGAGAGACTCTTGAAGAGTCCCTTGA 2467
Db 220997 TCCAAAGAAATGATGCTTTTGAACCTGTGGTGGAGAGACTCTTGAAGAGTCCCTTGA 220938
QY 2468 CTGCAAGAGATCCAAAGCAGTCCATTTGAGAGAGATCAGCCCTGGATTTCTTGGAG 2527
Db 220937 CTGCAAGAGATCCAAAGCAGTCCATTTGAGAGAGATCAGCCCTGGATTTCTTGGAG 220878
QY 2528 GAATGATGTAAAGCTGAAACTCCAGTACTTGGCCACCTGATCAGAAAGAGTGAATCAG 2587
Db 220877 GAATGATGTAAAGCTGAAACTCCAGTACTTGGCCACCTGATCAGAAAGAGTGAATCAG 220818
QY 2588 TGGAAAGACCTGATGCTGGAGAGGATTTGGGGCAGAGAGAGAGAGAGAGAGAGAGAGAG 2647
Db 220817 TGGAAAGACCTGATGCTGGAGAGGATTTGGGGCAGAGAGAGAGAGAGAGAGAGAGAGAG 220758
QY 2648 TGAGATGCTGGATGGCATCACTGATCTGATGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2707
Db 220757 TGAGATGCTGGATGGCATCACTGATCTCAATGAGCATGATCTGAGTGAATCTCCGGAGT 220698
QY 2708 TGGTGTATGAGCAGGGAGGCTGCTGCGCGGATTCATGGGTCACAAAGAGTGGACAC 2767
Db 220697 TGGTGTATGAGCAGGGAGGCTGCGATGCTGCGATTCGCGATTCGCGATTCGCGATTCGCGAT 220638
QY 2768 GACTGAGCAACTGAATGAACTGTAAGTGAACCTTAGTAGTTTATATTAATCT--- 2823

[illegible]

RESULT 3	
AC145008/c	
LOCUS	198916 bp DNA linear HTG 03-OCT-2003
DEFINITION	Bos taurus clone RP42-385L9, WORKING DRAFT SEQUENCE, 8 ordered pieces.
ACCESSION	AC145008
VERSION	AC145008.3 GI:37497133
KEYWORDS	HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE	Bos taurus (cow)
ORGANISM	Bos taurus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Ruminantia; Bos.
	1 (bases 1 to 198916)
REFERENCES	Antonellis, A., Avele, K., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Chu, G., Coleman, B., Coleman, H., Engle, J., Granite, S., Guan, X., Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Hu, P., Hurie, B., Ido, J.R., Karlins, E., Kwong, P., Lari, P., Larson, S., Lee-Lih, S.-Q., Legaspi, R., Maduro, Q.L., Maduro, V.B., Margulies, E.H., Masiello, C., Maskeri, B., McDowell, J., Mullikin, J.C., Paquirigan, C., Pearson, R., Portnoy, M.B., Prasad, A., Reddi-Dugue, N., Schandler, K., Schueler, M.G., Shah, K., Sison, C., Stantropop, S., Thomas, J.W., and Thomas, P.J., Tsiouri, V., Vogt, J.L., Wetherby, K.D., Young, A., and Green, E.D.

TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
-------	---------	-----------	---------	-------	---------	-----------	---------	-------	---------	---------

NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 198916)
Green, E.D.
Direct Submission
Submitted (30-MAY-2003) NIH Intramural
Government Circle, Gaithersburg, MD 208
2 (bases 1 to 198916)
Green, E.D.
Direct Submission
Submitted (03-OCT-2003) NIH Intramural
Government Circle, Gaithersburg, MD 208
On Oct 3, 2003 this sequence version r
----- Genome Center
Center: NIH Intramural Sequencing
Center code: NISC
Web site: <http://www.nisc.nih.gov>
Contact: nisc_zoon@nigrl.nih.gov
----- Project Information
Center project name: eeo
Center clone name: 385109

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

Summary statistics

Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 197481 bases at least Q40
Consensus quality: 197872 bases at least Q30
Consensus quality: 198094 bases at least Q20
Insert size: 153000; agarose- ϕ
Insert size: 198216; sum-of-contigs
Quality coverage: 15.52x in Q20 bases; agarose- ϕ
Quality coverage: 11.98x in Q20 bases; sum-of-contigs

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
*
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1 19088: contig of 19088 bp in length
* 19089 19188: gap of unknown length
* 19189 21334: contig of 2146 bp in length
* 21335 24434: gap of unknown length
* 24435 37376: contig of 15942 bp in length
* 37377 37476: gap of unknown length
* 37477 109937: contig of 72461 bp in length
* 109938 157979: gap of unknown length
* 157979 189815: contig of 47942 bp in length
* 189815 197980: gap of unknown length
* 197980 160129: contig of 2050 bp in length
* 160129 160130: gap of unknown length
* 160130 170533: contig of 10303 bp in length
* 170533 170633: gap of unknown length
* 170633 198915: contig of 28294 bp in length.

```

FEATURES	SOURCE
----------	--------

1. .198916

misc_feature		/mol_type="genomic DNA"	
		/db_xref="taxon:9913"	
		/clone="RP42-38519"	
		/clone_lib="RP42"	
		1..23685	
misc_feature		/note="clone overlaps with GenBank Accession Number AC14602 clone RP42-212N23 (center project name een)"	
		1..19088	
		/note="assembly_fragment"	
		clone_end:T7	
		vector_side:left	
misc_feature		19189..21334	
		/note="assembly_fragment"	
misc_feature		21435..37376	
		/note="assembly_fragment"	
misc_feature		37477..109937	
		/note="assembly_fragment"	
misc_feature		110038..157979	
		/note="assembly_fragment"	
misc_feature		158080..160129	
		/note="assembly_fragment"	
misc_feature		160230..170532	
		/note="assembly_fragment"	
misc_feature		170633..198916	
		/note="assembly_fragment"	
		clone_end:SP6	
		vector_side:right	
Query Match		60.3%; Score 2153.4; DB 1; Length 198916;	
Best Local Similarity		77.5%; Pred. No. 0;	
Matches 2630; Conservative		0; Mismatches 156; Indels 606; Gaps	8;
QY	1	GTCCAGGAGCGCGCAGTGGAGAGAGTACCTACCTCTGTCGACAGTAAAGGACGACGTAGC	60
Db	198496	GTCCAGGAGAGCGGTGTGAGGAGA-----TAACCTCTGTCGACAGTAAAGGACGACGTAGC	198442
QY	61	TGCCCTTTGTGGAGCGCGTAAAGAGATACCCCGCCCAAGGTAAAGAGAACCCCAAG	120
Db	198441	TGTCCTTTGTGGAGCGCGTGAAGAGATATCCAGCTCCAGGTAAAGAGAACCCCAAG	198382
QY	121	TAAGATGGTAGTGTGTGGAGGGCATCAGAGGGCAGACATCTGAACCATACACGCA	180
Db	198381	TAAATGGTAGTGTGTTCAGAGGGCATCAGAGGGCAGACATCTGAACCATACACGCA	198322
QY	181	GAATACTAGTCAATCTAATCACACTAGGACACAGCCCTGTCTTAAGCTCAATGAACCTAAG	240
Db	198321	GAATACTAGTCAATCTAATCACACTAGGACACAGCCCTGTCTTAAGCTCAATGAACCTAAG	198262
QY	241	CCATGCCCTGGGGCAACCCCAAGATGGCGAGGTCTGTTGGAGAGATCTGACAGAAATGTG	300
Db	198261	CCATGCCCTGGGGCAACCCCAAGATGGCGAGGTCTGTTGGAGAGATCTGACAGAAATGTG	198202
QY	301	GTCCACTGGAGAGAGGGAAT-GCAACCACTTCAGTATCTTCTGCTTGAGAACCCCATGAA	359
Db	198201	GTCCACTGGAGAGAGGGAATGCAACCACTTCAGTATCTTCTGCTTGAGAACCCCATGAA	198142
QY	360	CAGTATGAAAGGCAAAATGATAGGATCTGAAAGAGGAACTCCCGAGGTCAAGTGGTC	419
Db	198141	CAGTATGAAAGGCAAAATGATAGGATCTGAAAGAGGAACTCCCGAGGTCAAGTGGTC	198082
QY	420	CCCATATGCTACTGGAGATCAGTGGAGAAATACTCCAGAAAGAAATGAAGAGATGGAGCC	479
Db	198081	CCCATATGCTACTGGAGATCAGTGGAGAAATACTCCAGAAAGAAATGAAGAGATGGAGCC	198022
QY	480	AAAGCAAAAGAAATACCCAGCTGTGGATGTGACTGGTGATATAAGCAAGGTCCGATGTG	539
Db	198021	AAAGCAAAAGAAATACCCAGCTGTGGATGTGACTGGTGATATAAGCAAGGTCCGATGTG	197962
QY	540	TAAAGAGCATATTTGATAGGAACCTTGGAAATGTGAGGTCCATGAATCAAGGCAAAATGGA	599
Db	197961	TAAAGAGCATATTTGATAGGAACCTTGGAAATGTGAGGTCCATGAATCAAGGCAAAATGGA	197902
QY	600	AGTGGTCAAAACAGAGATGGCAAGAGTGAATGTCAACATTCAGGAATCAGCGNACTAA	659
Db	197901	AGTGGTCAAAACAGAGATGGCAAGAGTGAATGTCAACATTCAGGAATCAGCGNACTGAA	197842
QY	660	ATGGACTGGAATGGGTGAATTTAACTCAGATGACCATTTATCTACTACTCTGGGGCAGGA	719
Db	197841	ATGGACTGGAATGGGTGAATTTAACTCAGATGACCATTTATCTACTACTCTGGGGCAGGA	197782
QY	720	ATCCCTCAGAAGAAATGGAGTAGCCATCATGTCAACAAAAGAGTCCGAAATTCAGTACT	779
Db	197781	ATCCCTCAGAAGAAATGGAGTAGCCATCATGTCAACAAAAGAGTCTGAAATTCAGTACT	197722
QY	780	TGGATGCACTCTCAAAAACGACAGAAATGATCTCTGTTTGTTCGAAAGCAAAACCAATCAA	839
Db	197721	TGGATGCACTCTCAAAAACGACAGAAATGATCTCTGTTTGTTCGAAAGCAAAACCAATCAA	197662
QY	840	TATCAGAGTAATCCCAAGTCTATGCCCCCAACCAAGTAATCTGCTGAAGAGCTGGAAGTGAACG	899
Db	197661	TATCAGAGTAATCCCAAGTCTATGCCCCCAACCAAGTAATCTGCTGAAGAGCTGGAAGTGAACG	197602
QY	900	GTCTATGAAGACCTTACAAGACCTTTTAGAACTTAACCCCAAAAAGAGTGTCTTCTCAT	959
Db	197601	GTCTATGAAGACCTTACAAGACCTTTTAGAACTTAACCCCAAAAAGAGTGTCTTCTCAT	197542
QY	960	TATAGGGACCTGGAAATGCAAAAGTAGGAAGCAAGAAACACCTGGAGTAAACAGGCAATTT	1019
Db	197541	TATAGGGACCTGGAAATGCAAAAGTAGGAAGCAAGAAACACCTGGAGTAAACAGGCAATTT	197482
QY	1020	TGGCCCTTGGAAATACGGAATGAAGCAGGCAAGAACTAATAGAGTTTTCGCAAGAAATGC	1079
Db	197481	TGGCCCTTGGAAATACGGAATGAAGCAGGCAAGAACTAATAGAGTTTTCGCAAGAAATGC	197422
QY	1080	ACTGTCTATGCAAAACACCCCTCTTCCAAACAACAAGAAAGACTCTACATGGACATC	1139
Db	197421	ACTGTCTATGCAAAACACCCCTCTTCCAAACAACAAGAAAGACTCTACATGGACATC	197362
QY	1140	ACCAGATGGTCAACACCGAAATCAGATTGATTATTTCTTTGCGCCCAAGAGTGGAGAAG	1199
Db	197361	ACCAGATGGTCAACACCGAAATCAGATTGATTATTTCTTTGCGCCCAAGAGTGGAGAAG	197302
QY	1200	CTCTATACAGTCAGCAAAACCAAGCAGGAGCTTACTTGGCTCAGATCATGAATCTCT	1259
Db	197301	CTCTATACAGTCAGCAAAACCAAGCAGGAGCTTACTTGGCTCAGATCATGAATCTCT	197242
QY	1260	TATTGCCAAATTCAGACTTAAATTTGAAGAGTGGGAAACCACTAGACTACCTCAGGTA	1319
Db	197241	TATTGCCAAATTCAGACTTAAATTTGAAGAGTGGGAAACCACTAGACTACCTCAGGTA	197182
QY	1320	AGACTAAATCCAATCCCTTATGATTATACAGTGGAAAGTGAAGAAATAGATTAAAGGCGCT	1379
Db	197181	AGACTAAATCCAATCCCTTATGATTATACAGTGGAAAGTGAAGAAATAGATTAAAGGCGCT	197122
QY	1380	AGATCTGATACAGAGTACCTAATGAATGACTATGACAGAGGTTTCACTGACATTTACAGGA	1439
Db	197121	AGATCTGATACAGAGTACCTAATGAATGACTATGACAGAGGTTTCACTGACATTTACAGGA	197062
QY	1440	CACAGGATCGAGACCATCCCATGGAAGAAATGCAAAAAGCAAAATGGCTGTCTGG	1499
Db	197061	CACAGGATCGAGACCATCCCATGGAAGAAATGCAAAAAGCAAAATGGCTGTCTGG	197003
QY	1500	GGAGCCCTTCAAAATAGCTGTGAAAGAGAGAGTGAAGCAAGGAAAAAGGAAAG	1559
Db	197002	GGAGCCCTTCAAAATAGCTGTGAAAGAGAGAGTGAAGCAAGGAAAAAGGAAAG	196943
QY	1560	ATAAAGCATCTGAATCGAGATTCCAAA-----	1588
Db	196942	ATAAAGCATCTGAATCGAGATTCCAAAATAGCAAGAGATGAAGAACCTTCTT	196883
QY	1589	-----	1588
Db	196882	CAGCAAAATGCAAGAAATAGAGGAAACCAACAGAAATGGGAAAGACTAGGATCTCTT	196823
QY	1589	-----	1588

Db 196822 CAAGAAATAGAGGTACCAAGGACATTTTCATGCAAGATGAGCTCGATAAAGACAG 196763
Qy 1589 ----- 1588
Db 196762 AAATGATGACCTAACAGACAGAGATATTAGAAAGAGGTGCAAGAATACACAGA 196703
Qy 1589 ----- 1588
Db 196702 AGAACTGTACAAAAGATCTTCACGACCCAGATATCAGATGGTGTGATCACTCACT 196643
Qy 1589 ----- 1588
Db 196642 AGAGCCAGACATCCTGGAAATGTGAAGTCAAGTGGGCTTTAGAAAGCATCACTATGAAGAA 196583
Qy 1589 ----- 1588
Db 196582 AGCTAGTAGAGGTATGGAATTCAGATTGAGCTATTCCAAATCCTGAAAGATGATGCTGT 196523
Qy 1589 ----- 1588
Db 196522 GAAAGTGTGCTCACTCAATATGCCAGCAAAATTTGGAAAACTCAGAGTGGCCACAGAGCTG 196463
Qy 1589 ----- 1588
Db 196462 GAAAAGTCACTTTTCATTCCAATCCCAAGAAAGCAATGCCAAAGATGTTCAAACTA 196403
Qy 1589 ----- 1588
Db 196402 CCACCAATTTGCACTCATCTCACGCTAGTCAAGTAAATCTCCAAATTTCTCCAAAGCCAG 196343
Qy 1589 -----GAACTTCAGTTGTTCAAGCTGGTTTATAGAAAGTCAAG 1626
Db 196342 GCTTCAGCAATATGTGAATCTTCAAGTGTGTTCAAGCTGGTTTATAGAAAGCCAG 196283
Qy 1627 AGGACAGAGACCAATTCGCAACATCTCTGTATCATGGAAGAAAGCAAGAGTTCCTCA 1686
Db 196282 AGGACAGAGATCAAAATTCGCAACATCTCTGTATCATCAAAAGCAGAGAGTTCCTCA 196223
Qy 1687 GAAAAACATCTATTTCGCTTTTATTGACTATGCAAAAGCCTTTGACTGTGGGGTCACAA 1746
Db 196222 GAAAAACATCTATTTCGCTTTTATTGACTATGCTAAAGCCTTTGACTGTGGATCACAA 196163
Qy 1747 TAACTGTGAAATTTCTGAAGGATGGAATACAGACACCACTGACCTGACTCTTGAA 1806
Db 196162 TAACTGTGAAATTTCTGAAGAGATGGAATACAGACCACTGACTGCTCTTGAG 196103
Qy 1807 AAATTTGTATGAGTACAGGAGCAACAGTTAGAACTGGACATGGAACACAGAGCTGGTT 1866
Db 196102 AAATTTGTATGAGTACAGGAGCAATAGTTAGAACTGGACATGGAACACAGAGCTGGTT 196043
Qy 1867 CCAAGTAGGAAAGAGATGTCTAAGGCTGTATATTGTCTACCCGCTTGTATTAACTTCTA 1926
Db 196042 CCAATAGGAAAGAGATGTCTAAGGCTGTATATTGTCTACCCGCTTGTATTAACTTATC 195983
Qy 1927 TGCAGAG-ACATCATGAAACGCTGGGCTGGAGAGACCAAGCTGGAATCAAGATTGC 1985
Db 195982 TGCAGAGTACATCATAGAAACGCTGGACTGGAGAGAAACCAAGCTGGAATCAAGATTGC 195923
Qy 1986 CGGAGAGAAATAGCAATTAACCTCAGATATGCAATGATGATACCACTTATGCGAGAAAGTGA 2045
Db 195922 CGGAGAGAAATATCAATAACCTCAGATATGCAATGATGATACCACTTATGCGAGAAAGGGA 195863
Qy 2046 AGAGAACTAAAAGCCTCTTGATGAAGGTGAAGAGAGAGTGAAGAAAGTGGCTTAA 2105
Db 195862 AGAGAACTAAAAGCCTCTTGATGAAGAAAGTGAAGAGAGAGTGAAGAAAGTGGCTTAA 195803
Qy 2106 GCTCAACATTCAGAAAAGAGATCATGGCATCTGTGCTCCATCACTACTTCATGGGAAATAGA 2165
Db 195802 GCTCAACATTCAGAAAAGAGATCATGGCATCTGAGTCCCACTTCAATGCAATAGA 195743
Qy 2166 TGGGAAACAGTGGAAACAGTGTGAGCTTTATTTTTTTGGGGGCTCCAAATCACTGCA 2225
Db 195742 TGGGAAACAGTGGAAACAGTGTGAGCTTTATTTTTTTT- -GGGGCTCCAAATCACTACA 195685

Qy 2226 GATGGTGAAGTCAAGCCTGAAATTTAAAGACACACTTACTCTTGGAGAAAGTTA- -ACC 2283
Db 195684 GATGGTGAAGTCAAGCCTGAAATTTAAAGAGCCTTACTCTTGGAGAAAGTTATGACC 195625
Qy 2284 AACCTAGATAGCATATTGAAAGCAGAGACATTAACCTTGGCAACAAAGCCCATCTAGTC 2343
Db 195624 AACCTAGATAGCATATTGAAAGCAGAGACATTAACCTTGGCGACAAAGGTTTGTCTAGTC 195565
Qy 2344 AAGCTATGTTTTCAGTGGTTCATGTATGATGTGAGAGTGTGAGACTGTGAGAAAGCT 2403
Db 195564 AATGCTATGG- -TTTTCTGTGGTTCATGTATGATGTGAGAGTGTGAGACTGTGAGAAAGCT 195506
Qy 2404 GAGCAGCTGAAGAAATTTGATGCTTTTGAACCTGTGTGTTTGGAGAAAGCTTCTTGAAGTCCCT 2463
Db 195505 GAGCGCCAAAGAAATTTGATGCTTTTGAACCTGTGTGTTTGGAGAAAGCTATTGAGAGTCCCT 195446
Qy 2464 TGGACTGCAAGAGATCCAAACAGTCCATTTCTGAAGGAGATCAGCCTTGGGATTTCTTTG 2523
Db 195445 TGGACTGCAAGAGATCCAAACAGTCCATTTCTGAAGGAGATCAGCCTTGGGATTTCTTTG 195386
Qy 2524 GAAAGAAATGATGCTAAAGCTGAAACTCCAGTACTTTTGGCCACCTGATCAGAAAGCTGAC 2583
Db 195385 GAAAGAAATGATGCTAAAGCTGAAACTCCAGTACTTTTGGCCACCTCATGCAAGAGTTCAC 195326
Qy 2584 TCAGTGGAAAGACCTGATGCTGGAGGAGATTTGGGCGAGAGAGAGAGGAGGACAG 2643
Db 195325 TCATTGGAAAGACTCTGATGCTGGAGGAGATTTGGGCGAGAGAGAGAGGAGATGACAG 195266
Qy 2644 AGGATGAGATGCTGGATGGATCACTGACTCGATGAGACGTGAGTCTGGTCAACTCTCTG 2703
Db 195265 AGGATGAGATGCTGGATGGATCACTGACTCGATGAGACGTGAGTCTGAGTGAAGTCTGA 195206
Qy 2704 GAGTTGGTATGAGCAGAGAGGAGGCTGCTCTGGGCGATTCATGGGTCACAAAGAGTTGG 2763
Db 195205 GAGTTGGTATGAGCAGAGAGGAGGCTTGGGCTGCTGGATTCATGGGTCGCCAAGAGTTGG 195146
Qy 2764 ACAGCACTGAGCAACTGAACTGAACTGAACTG 2795
Db 195145 ACAGCACTGAGCAACTGATCTGATCTGATCTG 195114

RESULT 4
AC135839/c
LOCUS

DEFINITION Bos taurus clone RP42-2H17, WORKING DRAFT SEQUENCE, 6 ordered
pieces.
ACCESSION AC135839
VERSION AC135839.2 GI:24942893
KEYWORDS HTG; HTGS; PHASE2; HTGS_DRAFT.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus

REFERENCE
AUTHORS

1 (bases 1 to 179015)
Akhtar N., Antonellis A., Ayele, K., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Cargill, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J.,
Raghani, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, S.,
Laric, P., Lee-Lin, S.-Q., Legaspi, R., Maduro, Q.L., Maduro, V.B.,
Margulies, E.H., Mastello, C., Maskeri, B., McDowell, J.,
Pagnirigan, C., Pearson, R., Portnoy, M.E., Prasad, A.,
Reddy-Dugue, N., Schandler, K., Schueler, M.G., Sison, C.,
Statrapiop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L.,
Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 179015)
Green, E.D.

TITLE
JOURNAL
REFERENCE

Direct Submission
Submitted (23-OCT-2002) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA

REFERENCE 3 (bases 1 to 179015)
 AUTHORS Green, E.D.
 TITLE Direct Submission
 JOURNAL Submitted (13-NOV-2002) NIH Intramural Sequencing Center, 8717
 Government Circle, Gaithersburg, MD 20877, USA
 COMMENT On Nov 13, 2002 this sequence version replaced gi:24270702.
 ----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: <http://www.nisc.nih.gov>
 Contact: nisc_zoonhgri.nih.gov
 ----- Project Information
 Center project name: cra
 Center clone name: 002H17

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics
 Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.950319
 Consensus quality: 177630 bases at least Q30
 Consensus quality: 178036 bases at least Q40
 Consensus quality: 178350 bases at least Q20
 Insert size: 158000; agarose-fp
 Insert size: 178515; sum-of-contigs
 Quality coverage: 10.16x in Q20 bases; agarose-fp
 Quality coverage: 8.99x in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 6 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.

* This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

* 1 94998: contig of 94998 bp in length
 * 94999 95098: gap of unknown length
 * 95099 100734: contig of 5636 bp in length
 * 100735 100835: gap of unknown length
 * 100835 108201: contig of 7367 bp in length
 * 108201 108302: gap of unknown length
 * 108302 135923: contig of 27622 bp in length
 * 135923 136023: gap of unknown length
 * 136023 166723: contig of 30700 bp in length
 * 166723 166824: gap of unknown length
 * 166824 179015: contig of 12192 bp in length.

FEATURES
 Source
 1. 179015
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9913"
 /clone="RP42-2H17"
 /clone_lib="RP42"

1. 94998
 /note="assembly_fragment
 clone_end:17
 vector_side:left"

misc_feature
 1. 57332
 /note="clone overlaps with GenBank Accession Number
 AC136201 clone RP42-131E20 (center project name cj2)"

misc_feature 95099..100734
 /note="assembly_fragment"
 misc_feature 100835..108201
 /note="assembly_fragment"
 misc_feature 108302..135923
 /note="assembly_fragment"
 misc_feature 136024..166723
 /note="assembly_fragment"
 misc_feature 166824..179015
 /note="assembly_fragment
 clone_end:SP6
 vector_side:right"
 Query Match 60.1%; Score 2147.7; DB 1; Length 179015;
 Best Local Similarity 77.6%; Pred. No. 0;
 Matches 2633; Conservative 0; Mismatches 153; Indels 607; Gaps 9;
 QY 1 GTCAGGAAGGGCGGCGAGTACCTGAGAGGAGTACCTGCTCCAGGTAAAGGAGCAGTAGC 60
 Db 146671 GTCAGGAAGGGCGGCGGTAGGAGA----TACCCCTCGTCCAGGTAAAGGAGCAGCGC 146617
 QY 61 TCGGCTTTCTGGAGCAGCCGTAAAGAGATACCCACGCCCAAGG-TAAGAGAAACCCAA 119
 Db 146616 TCGGCTTTCTGGAGCAGCTCTGAAGAGATAGCCACACCCCAAGGTTAAGAGAAACCCAA 146557
 QY 120 GTAAGATGTTAGTGTCTGTGAGAGGGCATCAGAGGGCAGACATACTGAAACCATACACGC 179
 Db 146556 GTAAGATGTTAGTGTCTGCAAGAGGGCATCCGAGGCGAGATGCATGAAACCATACTCAC 146497
 QY 180 AGAAACTAGTCAATCTAATCAGTAGGACACAGCCCTTGTCTTAATCAATGAAACTAA 239
 Db 146496 AGAAAGCTAGTCAATCTAATCAGTAGGACACAGCGGTATCTTAATCAATGAAACTAA 146437
 QY 240 GCATGCCCGTGGGGCAACCCAGATGGCGAGGTCTGTTGGAGAGATCTTCACAGAAATGT 299
 Db 146436 GCATGCCCGTGGGGCAACCCAGAGCGGCGGTCTGTTGGAGAGATCTTCACAGAAATGT 146377
 QY 300 GGTCCATGGAAGAGGGAAAT-GCAAACCATCTTCACTTCTTTCCTTTCAGAACCCCATGA 358
 Db 146376 GGGCCACCGGAGAGAGAAATGGCAACCCATCTCAGTATCTTTCCTTTCAGAACCCCATGA 146317
 QY 359 ACAGTATCAAGAGCAGAAATGATAGTACTGAAAGAGGAGTCCCGAGGTCACTAGGTG 418
 Db 146316 ACAGTATCAAGAGCAGAAATGATAGTACTGAAAGAGGAGTCCCGAGGTCACTAGGTG 146257
 QY 419 CCCCATGTCTACTGGAGATCAGTGGAGAAATAACTCCAGAGAAATGAAGAGATGGAGC 478
 Db 146256 CCCCATGTCTACTGGAGATCAGTGGAGAAATAACTCCAGAGAAATGAAGAGATGGAGC 146197
 QY 479 CAAGACAAAAGATACCAGCTGTGGATGTGATGTGATATAGCAAGTCCGATGCT 538
 Db 146196 CAAGAGAAAACAATACCCAGCTGTGGATGTGATGTGATATAGCAAGTCCGATGCT 146137
 QY 539 GTAAAGAGCAATATTGCATAGGAACCTGGAATGTGAGTCCATCAATCAAGGCAAAATGG 598
 Db 146136 GTAAAGAGCAATATTGCATAGGAACCTGGAATGTGAGTCCATCAATCAAGGCAAAATGG 146077
 QY 599 AAGTGGTCAACCAAGAGATGCAAGAGTGAATGTCAATCTTAGGAATCAGCGAACTAA 658
 Db 146076 AAGTGGTCAACCAAGAGATGCAAGAGTGAATGTCAATCTTAGGAATCAGCGAACTAA 146017
 QY 659 AATGACTGGAAATGGGTGAATTTAACTCAGATGACCATTTATCTACTCTGCGGGCAGG 718
 Db 146016 AATGACTGGAAATGGGTGAATTTAACTCAGATGACCATTTATCTACTCTGCGGGCAGG 145957
 QY 719 AATCCCTCAGAGAAATGGAGTAGCCATCATGTGTCACAAAAGAGTCCGAAATGCACTAC 778
 Db 145956 AATCCCTCAGAGAAATGGAGTAGCCATCATGTGTCACAAAAGAGTCCGAAATGCACTAC 145900
 QY 779 TTGGATCGAGTCTCAAAAACGACAGAGATGATCTCTGTTTCTTCCAGGCAAAACCATTCAC 838
 Db 145899 TTGGATCGAGTCTCAAAAACGACAGAGATGATCTCTGTTTCTTCCAGGCAAAACCATTCAC 145840

QY	839	ATATCACAGTAATCCAAAGTCTATGCCCCCAACCAAGTAATGCTCAGAAAGCTGAAGTTGAAC	898	Db	144760	GCTAGTGGAGGTGATAGAAATTCAGTTGAGCTATTCCAAATCCTGAAAGATGATGCTGTG	144701
Db	145839	ATATCACAGTAATCCAAAGTCTGTCGCCCAACCAAGTAATGCTCAGAAAGCTGAAGTTGAAC	145780	QY	1589	-----	1588
QY	899	GGTCTATCAAGACCTACAAAGCTTTTGAAGACTAACACCCCAAAAGAGATGCTCTTCTCA	958	Db	144700	AAAGTGTGCACTCAATATGCCAGCAAAATTTGGAAAACCTCAGCAGTGACCACAGGACTGG	144641
Db	145779	GGTCTATCAAGACCTACAAAGCTTTTGAAGACTAACACCCCAAAAGAGATGCTCTTCTCA	145720	QY	1589	-----	1588
QY	959	TTATAGGGGACTGGAATGCAAAAGTAGGAAGCAAAAGAACACCTGAGTAGCAGCAAAAT	1018	Db	144640	AAAGSTCAGTTTTTCATTCCAATCCCAAGAAAGCAATGCCAAAGAAATGCTCAAACTAC	144581
Db	145719	TTATAGGGGACTGGAATGCAAAAGTAGGAAGCAAAAGAACACCTGAGTAGCAGCAAAAT	145660	QY	1589	-----	1588
QY	1019	TTGGCCCTTGAATACGGAATGAAGCAGGCAAGACTAATAGATTTTGGCCAAAGAAAATG	1078	Db	144580	CACACAATTCGACTCATCTCACATGCTPAATAAAGTAATACTCAAAATTCCTCAAGCCAGG	144521
Db	145659	TTGGCCCTTGAATACGGAATGAAGCAGGCAAGACTAATAGATTTTGGCCAAAGAAAATG	145600	QY	1589	-----	1588
QY	1079	CACTGCTATAGCAAAACACCCCTTTCACAAACAACAAGAAAGACTCTACATGAGACAT	1138	Db	144520	CTTCAGCAATACGTGAACCGTGAACCTTCAGATGTTCCAGATGTTTTCAGAAAAGGCAGA	144461
Db	145599	CACTGCTATAGCAAAACACCCCTTTCACAAACAACAAGAAAGACTCTACATGAGACAT	145540	QY	1628	GGAAACAGAGACCAAAATTCGCAACATCCTCTGTATCATGTGAAAAGCAAGAGATTCCAG	1687
QY	1139	CACCAGATGGTCACACCGGAATCAGATTGATTATTTCTTTGACGCCAAAGATGGAGAA	1198	Db	144460	GGGACCGAGAGACCAAAATTCGCAACATCCTGCTGGATCATGAAAAAAGCAAGAGATTCCAG	144401
Db	145539	CACCAGATGGTCACACCGGAATCAGATTGATTATTTCTTTGACGCCAAAGATGGAGAA	145480	QY	1588	AAAAACATCTATTTCTGCTTTTATGACTATGCAAAAGCCCTTTGACTGTGGGGGTCACAAT	1747
QY	1199	GCTCTATACAGTCAGCAAAACCAAGACCAAGAGCTTTACTGTGGCTCAGATCATGAATGCC	1258	Db	144400	AAAGACATCTATTTCTGCTTTTATGCTTTTGGCGATGCCAAAGCCCTTTAAGTGTGGGTCACAAT	144341
Db	145479	GCTCTATACAGTCAC - CAACCAAGACGGGAGCTGACTGTGGCTCAGATCATGAATGCC	145421	QY	1748	AACTGTGGAAAATTCGAAAGGATGGGAATACCCAGACCACTGACCTGACTCTTGAAA	1807
QY	1259	TTATGGCCAAATTCAGACTTTAATTTGAAGAAAGTAGGAAACCACTAGATCACTCAGGT	1318	Db	144340	AACTGTGGACGATTCGAAAGACATGGGAATACCCAGACCACTGACCTGCTCTTGAGA	144281
Db	145420	TTATGGCCAAATTCAGACTTTAATTTGAAGAAAGTAGGAAACCACTAGATCACTCAGGT	145361	QY	1808	AATTTGTATGAGTTCAGGAAGCAACAGTTAGACTGGACATGGAACAACAAGACTGGTTC	1867
QY	1319	AAGACCTAAATCCAAATCCCTTATGATTATACAGTGGAAAGTAGAAATAGATTAAAGGGCC	1378	Db	144280	AATCTGTATGAGTTCAGGAAGCAACAGTTAGACTGGACATGGAACAACAAGACTGGTTC	144221
Db	145360	ATGACCTAAATCCAAATCCCTTATGATTATACAGTGGAAAGTAGAAATAGATTAAAGGGCC	145301	QY	1868	CAAGTAGGAAAAGGAGTATGTCAGAGCTGTATTTGTTCACCCGGCTGTGTTTAACTCTTAT	1927
QY	1379	TAGATCTGATAGACAGAGTACCTTAATGAATATGATGACAGAGTTATGACATTTGACAGG	1438	Db	144220	CAATAGACAGAGGAGTACGTCAGGCTGTATTTGTTCACCCCTGCTTTTAACTTATAT	144161
Db	145300	TACATCTGATAGATAGTGCCTGATGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT	145241	QY	1928	GCAGA - GACATCATGAGAAAACGCTGGGCTGGAAGACACAGCTGAGTCAAGATTGCC	1986
QY	1439	AGACGGGATCGAGACATCCCATCGAAAGAAAGTGCAAAAGCAAAATGCTGCTCTG	1498	Db	144160	GCAGATTACATCATGAGAAAACGCTGGACTGGAAGAAAACACAAGCTGGAATCAAGATTGCC	144101
Db	145240	AGACGGGATCGAGACATCCCATCGAAAGAAAGTGCAAAAGCAAAATGCTGCTCTG	145181	QY	1987	GGGAGAAATAGCAATAACCTCAGATATGAGATATGACATGATGACCCCTTATGGCAAGAAAGTAA	2046
QY	1499	GGGAGGCTTTACAAATAGCTGTGAAAAGAGAGAGAGTGAAGCAAGCAAGCAAGCAAGCAAG	1558	Db	144100	GGGAGAAATATCATTAACCTCAGATATGAGATGACATGACCCCTTATGGCAAGAAAGTAA	144041
Db	145180	GGGAGGCTTTACAAATAGCTGTGAAAAGAGAGAGAGTGAAGCAAGCAAGCAAGCAAGCAAG	145121	QY	2047	GAGGAACTAAAAGGCTCTTGATGAAGGTGAAAGGAGAGAGTGAAGAAAGTGGCTTAAAG	2106
QY	1559	GATTAAGAGATCTGAATGAGAGTTCCAA -	-----	Db	144040	GAGGAACTAAAAGGCTCTTGATGAAGGTGAAAGGAGAGAGTGAAGAAAGTGGCTTAAAG	143981
Db	145120	GATTAAGAGATCTGAATGAGAGTTCCAAAGATTCCAAGATAGCAAGAGATAGCAAGAGCTTCC	145061	QY	2107	CTCAACATTCAGAAAACGAGATCATGGCATCTGTCCTCCATCCTCATGGGAAATAGAT	2166
QY	1589	-----	1588	Db	143980	CTCGAATTCAGAAAACGAGATCATGGCATCTGTCCTCCATCCTCATGGGAAATAGAT	143921
Db	145060	TCAGCGATCAATGCAAGAAATAGAGGAAACCAACAGATGGTGAAGACTAGAGATCTCT	145001	QY	2167	GGGGAAAACAGTGGAAAACAGTGTACACTTTATTTTTGGGGGGCTCCAAAATCACTGCAG	2226
QY	1589	-----	1588	Db	143920	GGGGAAAACAGTGGAAAACAGTGTGCTGACTTTA - TTTTGGGGGGCTCCAAAATCACTGCAG	143863
Db	145000	TCAGAAAATTTAGAGATACCAAGGGGACATTTTCATGCAAGAGATGGGCTTGATAAGGGCA	144941	QY	2227	ATGGTGAATGACAGCCATGAAAATTAAGACACTTACTCTCTTCTTGGGAAAGAAAGTTA - ACCA	2284
QY	1589	-----	1588	Db	143862	ATGGTGAATGACAGCCATGAAAATTAAGACACTTACTCTCTTCTTGGGAAAGAAAGTTAAGCA	143803
Db	144940	GAAATGAGGGACCTTAACAGACGACAGAGATATTAAGAGAGATGGCAAGATACACAG	144881	QY	2285	ACCTAGATAGCATATTTGAAAACGAGAGATTTACCTTGCCAAACAAAGCCCCATCTAGTCA	2344
QY	1589	-----	1588	Db	143802	AGCTAGATAGCATATTTGAAAACGAGAGATTTACCTTGCCAAACAAAGGGTCTCTCTAGTCA	143743
Db	144880	AAGAACTGTACAAAAGATCTTTCACGACCCAGATATACGATGGTGTGATCACTCATC	144821	QY	2345	AGGCTATGTTTTTCCAGTGGTTCATGTATGGATGTGAGAGTTGGACTGTGGAAGAAAGCTG	2404
QY	1589	-----	1588	Db	143742	AGGCTATGTTTTTCTGTGGTTCATGTATGGATGTGAGAGTTGGACTGTGGAAGAAAGCTG	143683
Db	144820	TAGAGCAGACATCTGGAATGTGAAGTCAAGTGGGCCCTTAGAAAGCATCACTACAAACAA	144761	QY	2405	AGCACTGAAGAAATTCATGCTTTTTTGAAGTGTGAGTGTGGAAGAGACTCTTTGAGAGTCCCTT	2464
QY	1589	-----	1588				

misc_feature	/note="assembly_fragment" 20978..30802	QY	599	AAGTGGTCAAAAGAGATGGCAAGTGAATGTCAACATTTCTAGGAATCAGCGAATAA	658
misc_feature	/note="assembly_fragment" 30903..55941	Db	66249	AAGTGGTCAAAAGAGATGGCAAGTGAATGTCAACATTTCTAGGAATCAGCGAATAA	66190
misc_feature	/note="assembly_fragment" 56042..86898	QY	659	AATGGACTGGAAATGGGTGAATTTAACTCAGATGACCATATATATCTACTACGCGGAGG	718
misc_feature	/note="assembly_fragment" 86999..123280	Db	66189	AATGGACTGGAAATGGGTGAATTTAACTCAGATGACCATATATATCTACTACGCGGAGG	66130
misc_feature	/note="assembly_fragment" 123381..158900	QY	719	AATCCCTCAGAAAGAAATGGAGTAGCCATCATGTGCTCAACAAAGAGTCCGAAATCAGATAC	778
misc_feature	/note="assembly_fragment" 159001..171357	Db	66129	AATCCCTCAGAAAGAAATGGAGTAGCCATCATGTGCTCAACAAAGAGTCCGAAATCAGATAC	66073
misc_feature	/note="assembly_fragment" 171458..178118	QY	779	TTGATGAGTCTCAAAAACACAGAAATGATCTCTGTTGTTTCCAGGCAAAACATTCAC	838
misc_feature	/note="assembly_fragment" 178219..183713	Db	66072	TTGATGCAATCTCAAAAATGACAGAATGATCTCTGTTGTTTCCAGGCAAAACATTCAC	66013
misc_feature	/note="assembly_fragment" 183814..231842	QY	839	ATATCACAGTAATCCAACTGTCATGCCCAACAGTAAATGCTGGAAGAGTCAAGTTGAAC	898
misc_feature	/note="assembly_fragment" 231943..237141	Db	66012	ATATCACAGTAATCCAACTGTCATGCCCAACAGTAAATGCTGGAAGAGTCAAGTTGAAC	65953
misc_feature	/note="assembly_fragment" 237242..241923	QY	899	GGTCTATGAAGACTACAGACCTTTTGAAGAACTAAACCCCAAAAGATGTCTCTTCTCA	958
misc_feature	/note="assembly_fragment" clone_end:17	Db	65952	GGTCTATGAAGACTACAGACCTTTTGAAGAACTAAACCCCAAAAGATGTCTCTTCTCA	65893
misc_feature	vector_side:right 237289..241923	QY	959	TTATAGGGGACTGGAAATCGAAAGTGAAGCAAGAAACACCTGGAGTAACAGCAAAAT	1018
misc_feature	AC136880 clone RP42-70M11 (center project name crd)"	Db	65892	TTATAGGGGACTGGAAATCGAAAGTGAAGCAAGAAACACCTGGAGTAACAGCAAAAT	65833
Query Match	60.1%; Score 2147.7; DB 1; Length 241923;	QY	1019	TTGGCTTTGGAAATACGGAAATGAAGCAGGGCAAGAACTAATAGAGTTTTCACCAAGAAATG	1078
Best Local Similarity	77.6%; Pred. No. 0;	Db	65832	TTGGCTTTGGAAATACGGAAATGAAGCAGGGCAAGAACTAATAGAGTTTTCACCAAGAAATG	65773
Matches 2633; Conservative 0; Mismatches 153; Indels 607; Gaps 9;		QY	1079	CACTGGTCATGAGCAAAACACCTCTTCAACAAACACAGAGAAAGACTCTACACATGGACAT	1138
1	GTGAGGAAGGCGCGAGTGAAGGAGGTACCTACCTCGTCCAAAGTAAGGACGATAGC	Db	65772	CACTGGTCATGAGCAAAACACCTCTTCAACAAACACAGAGAAAGACTCTACACATGGACAT	65713
66844	GTGAGGAAGGCGCGGTGAGGAGA-----TACCCTCGTCCAAAGTAAGGACGACGGC	QY	1139	CACCAAGTGGTCAACACCGAAATCAGATTGATTAATTTCTTTCAGACCAAGAGATGGAGAA	1198
61	TGGCTTTGCTGAGCAGCGCGTAAAGAGATACCCACGCGCCACAGG-TAAGAGAAACCAA	Db	65712	CACCAAGTGGTCAACACCGAAATCAGATTGATTAATTTCTTTCAGACCAAGAGATGGAGAA	65653
66789	TGGCTTTGCTGAGCAGCGCTCTGAAGAGATACCCACACCCAAAGGTTAAGAGAAACCCAA	QY	1199	GCTCTATACGTACGCAAAAAACAGACCAGAGCTTACTGTGGCTCAGATCATGAACCTCC	1258
120	GTAAAGTGGTGGTGTGTGAGAGGGCATCAGAGGGCAGACATCTCAAAACCATACACGC	Db	65652	GCTCTATACGTACGCAAAAAACAGACCAGAGCTTACTGTGGCTCAGATCATGAACCTCC	65594
66729	GTAAAGTGGTGGTGTGTGAGAGGGCATCAGAGGGCAGACATCTCAAAACCATACACGC	QY	1259	TTATTTGCCAAATTCAGACTTAAATTTGAGAAAGTAGGGAACCACTAGATCACTCAGGT	1318
180	AGAAAACTAGTCAATTAATCACTAGGACCAACAGCCTTGTCTAATCAATGAACCTAA	Db	65593	TTATTTGCCAAATTCAGACTTAAATTTGAGAAAGTAGGGAACCACTAGATCACTCAGGT	65534
66669	AGAAAGCTAGTCAATTAATCACTAGGACCAACAGCCTTGTCTAATCAATGAACCTAA	QY	1319	AAAGACTTAAATCCCAATCCCTTATGATTTATACAGTGGAAAGTGAAGAAATAGATTTAAGGGCC	1378
240	GCCATGCCGTGGGCAACCAAGATGGCGAGTCTATGGTGGAGAGATCTGACAGAAATGT	Db	65533	AAAGACTTAAATCCCAATCCCTTATGATTTATACAGTGGAAAGTGAAGAAATAGATTTAAGGGCC	65474
66609	GCCATGCCGTGGGCAACCAAGATGGCGAGTCTATGGTGGAGAGATCTGACAGAAATGT	QY	1379	TAGATCTGATGACAGAGTACCTAATGAACCTATGGACAGAGGTTTCACTGACATTTGTACAGG	1438
300	GGTCCACTGGAGAGGGAAT-GCAAAACACTTCAGTATTTCTGCTTGGAGAACCCCATGA	Db	65473	TACATCTGATGATAGAGTGCCTGATGAAGTATGGAATGAGGTTTGTGACACTGTACAGG	65414
66549	GCGCCACCGGAGAGAGAAATGGCAAAACCACTTCAGTATTTCTTGCCTTGGAAACCCCATGA	QY	1439	AGACAGGGATCAGACCACTCCCATGGAAAGAAATGCAAAAAAGCAAAATGGCTGTCTG	1498
359	ACAGTATCAAAAGGCAAAATGATAGGATCTGAAAGAGGAACCTCCCAAGTCAAGTGGT	Db	65413	AGACAGGGATCAGACCACTCCCATGGAAAGAAATGCAAAAAAGCAAAATGGCTGTCTG	65354
66489	ACAGTATCAAAAGGCAAAATGATAGGATCTGAAAGAGGAACCTCCCAAGTCAAGTGGT	QY	1499	GGGAGGCTTCAAAATAGCTGTGAAAGAGAGAGTCAAAAGCAAGCAAAAGGAGAA	1558
419	CCCATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGAAATGAAGAGATGGAGC	Db	65353	GGGAGGCTTCAAAATAGCTGTGAAAGAGAGAGTCAAAAGCAAGCAAAAGGAGAA	65294
66429	CCCAATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGAAATGAAGAGATGGAGC	QY	1559	GATTAAGCAATCTGAATGCAGAGTTCCAAAGNATAGCAAGAGAGATAGAAGCGCTTCC	1588
479	CAAGCAAAAAGAAATACCCAGCTGTGGATGTGATGTGATTAAGCAAGGTCGATGCT	Db	65293	GATTAAGCAATCTGAATGCAGAGTTCCAAAGNATAGCAAGAGAGATAGAAGCGCTTCC	65234
66369	CAAGAGAAAACATACCCAGCTGTGGATGTGATGTGATTAAGCAAGGTCGATGCT	QY	1589	-----	1588
539	GTAAGAGCAATATTGATAGGACCTGGAAATGTCAGTCCATGAATCAAGGCAAAATTGG	Db	65233	TCAGCGATCAATGCAAGAAATAGAGGAAACCAACAGAAATGGTAAAGACTAGAGATCTCT	65174
66309	GTAAGAGCAATATTGATAGGACCTGGAAATGTCAGTCCATGAATCAAGGCAAAATTGG	QY	1589	-----	1588

Db	65173	TCAAGAAAAATTAGAGATACCAAGGGGACATTTTCATGCAAGAAGATGGCTTTCATAAAGGCA	65114
Qy	1589	-----	1588
Db	65113	GAATGCTAGGACCTAACAGACGACGAGATATTAAAGAGAGATGGCAAGAATACACAG	65054
Qy	1589	-----	1588
Db	65053	AAGAACTGTACAAAAAAGATCTTCACGACCCAGATAATCAGCATGGTGTGATCACTCATC	64994
Qy	1589	-----	1588
Db	64993	TAGAGCAGACATCTGGAACTGTGAAGTCAAGTGGGCCCTTAGAAGCATCCTACAAACAAA	64934
Qy	1589	-----	1588
Db	64933	GCTAGTGGAGGTGATAGAATTCAGTTGAGCTATTCCAAATCCTGAAAGATGATGCTGTG	64874
Qy	1589	-----	1588
Db	64873	AAAGTCTGCATCAATATGTCAGCAAAATTTGGAAACTCAGCATGACCACGAGACTGG	64814
Qy	1589	-----	1588
Db	64813	AAAAGTCACTTTTCATTCCAATCCCAAGAAAGGCAATGCCAAAGATGCTCAAACCTAC	64754
Qy	1589	-----	1588
Db	64753	CACACAATTGCATCTCATCTCACATGCTAATAAAGTAATACTCAAAATTTCTCCAGCCAGG	64694
Qy	1589	-----GAACTTCCAGTTGTTTCAAGCTGGTTTAGAAAGTTCAG	1627
Db	64693	CTTCAGCAATACGTGAACCGTGAACCTCCAGATGTTCAAGCTGGTTTAGAAAGGCGAGA	64634
Qy	1628	GGAAACGAGACCAAAATTTGCCAATCTCTGTATCATGGAAGGAGAGAGAGTTCCAG	1687
Db	64633	GGGACCGGAGACAAAATTTGCCAATCCGCTGGATCATGGAAGGAGAGAGAGTTCCAG	64574
Qy	1688	AAAAACATCTATTTCTGCTTTATTGACTATGCAAAAGCCTTTGACTGTGGGGGTACAAAT	1747
Db	64573	AAAAGCATCTATTTCTGCTTTATTGGCGATGCCAAAGCCTTTAAGTGTGTGGGTCACAAT	64514
Qy	1748	AAACTGTGGAATAATTCTGAAAGGATGGGAATACCGAGCCACTGACCTGACTCTTTGAAA	1807
Db	64513	AAACTGTGGACGATTTCTGAAAGACATGGGAATACCGAGCCACTGACCTGCTTTGAGA	64454
Qy	1808	AAATTGTATCGAGTTCAGGAGCAACAGTTAGACTGGACATGGAAACACACAGACTGGTTC	1867
Db	64453	AACTGTGTATCGAGTCAGGAAGCAACAGTTAGACTGGACATGGAAACACACAGACTGGTTC	64394
Qy	1868	CAAGTAGGAAAAAGGATGTCTCAAGGCTGTATATTGTCAACCGGCTTGTTTAACTTCTAT	1927
Db	64393	CAANTAGGCAAGGATACGTCAAGGCTGTATTGTCAACCTGCTTATTTAATCTTAT	64334
Qy	1928	GCAGA-GACATCATGAGAAACCGCTGGGCTGGGAAGAGCAACAGCTGGAAATCAAGATTGCC	1986
Db	64333	GCAGATTACATCATGAGAAACCGCTGGACCTGGGAAGAAACACAGCTGGGAATCAAGATTGCC	64274
Qy	1987	GGGAGAAATAGCAATAACCTTCAGATATGCAGATGATACCACCTTTATGGCAGAAAGTGAA	2046
Db	64273	GGGAGAAATATCAATAACCTCAGATATGCAGATGACCAACCTTTATGGCAGAAAGTGAA	64214
Qy	2047	GAGGAATCAAAAAGCCTCTTGATGAAGGTGAAGAGGAGAGTGAAAAAGTTGGCTTTAAAG	2106
Db	64213	GAGGAATCAAAAAGCCTTTTGATGAAAGTGAAGAGGAGAGTGAAAAAGTTGGCTTTAAAG	64154
Qy	2107	CTCAACATTCAGAAACGAGAGATCATGGCATCTGCTGCCATCTCACTTCATGGGAATAGAT	2166
Db	64153	CTCGAATTCAGAAAAACGAGAGATCATGGCATCCCGTCCCATCTCACTTCATGGGAATAGAT	64094
Qy	2167	GGGGAAACAGTGGAAAAACAGTGTCAAGCTTTATTTTTTGGGGGGCTTCCAAAAATCACTGCAG	2226

Db	64093	GGGGAAACAGTGGAAAGAGAGTGCGCTGACTTTTA--TTTGTGGGGGGCTCCAAAATATCACTGCGAG	64093
QY	2227	ATGGTGACTCGACCCATGAAATTTAAAGACACACTTACTCTCTTGGAGAAAAGTTA--ACCA	2284
Db	64035	ATGGTGACTCGACCCATGAAATTTAAAGACACACTTACTCTCTTGGAGAAAAGTTATGACCA	63976
QY	2285	ACCTAGATAGCATATTGAAAAGCAGAGACATTTACCTTGGCCACAAAGCCCATCTAGTCA	2344
Db	63975	AGCTAGATAGCATATTGAAAAGCAGAGACATTTACTTTGGAAACAAAGTCTCTCTAGTCA	63916
QY	2345	AGGCTATGGTTTTTCCAGATGGTCAATGATGGATGTGAGAGTTTGGACTGTGAAAGAAAGCTG	2404
Db	63915	AGGCTATGGTTTTTCCGTGGTCAATGATGGATGTGAGAGTTTGGACTGTGAAAGAAAGCTG	63856
QY	2405	AGCACTGAAGAAATGTATGCTTTTGAACCTGGTGTGGAGAGACATCTTTGAGAGTCCCTT	2464
Db	63855	AGCGCTGAAGAAATGTATGCTTTTGAACCTGGTGTGGAGAGACATCTTTAAGAGTCCCTT	63796
QY	2465	GGACTGCAAGGAGATCCAAACAGTCCATTTCTGAAGAGAGATCAGCCCTGGGATTTCTTTGG	25224
Db	63795	GGACTGCAAGGAGATCCAAACAGTCCATTTCTGAAGAGAGATCAGCCCTGGGATTTCTTTGG	63736
QY	2525	AAGGAATGATGCTAAAGCTGAAACTCCACTACTTTTGGCCACTGATCAGAAGAGTGTACT	2584
Db	63735	AAGGAATGATGCTGAAAGCTGAAACTCCACTACTTTTGGCCACTGATCAGAAGAGTGTACT	63676
QY	2585	CATCTGNAAGACCCCTGATGCTGGGAGGATTTGGGGCAGGAGGAGAGGGGACGACAGA	2644
Db	63675	CATTGGAAAAGACTGTGATGCTGGGAGGATTTGGGGCAGGAGGAGAGGGGACGACAGA	63616
QY	2645	GGATGAGATGGCTGGATGGCATCACTGACTCGATGGACCTGAGTCTGGTGTGAACCTCTGG	2704
Db	63615	GGATGAGTGGCTGGATGGCATCACTGACTCGATGGACCTGAGTCTGGTGTGAACCTCTGG	63556
QY	2705	AGTTGGTGTGAGCAGGAGGCGTGTCTCGCGCGATTCATGGGCTCAGAAAGTGTGA	2764
Db	63555	AGTTGGTGTGAGCAGGAGGCGTGTCTCGCGCGATTCATGGGCTCAGAAAGTGTGA	63496
QY	2765	CACGACTGAGCACTCAACTGCAACTGAACTGTA	2797
Db	63495	CACGACTGAGCACTGCAACTGCAACTGTA	63463
RESULT 6			
AC091660	AC091660	173068 bp	DNA linear
DEFINITION	Bos taurus clone RP42-354B6,	complete sequence.	
ACCESSION	AC091660		
VERSION	AC091660.2	GI:15011695	
KEYWORDS	HTG.		
SOURCE	Bos taurus (cow)		
ORGANISM	Bos taurus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.		
AUTHORS	1 (bases 1 to 173068)		
	Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Ho, S.-L., Idol, J.R., Karlins, E., Lee-lin, S.-Q., Legaspi, R., Lim, M., Maduro, Q.L., Maduro, V.B., Masiallo, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Prasad, A., Shevchenko, Y., Snyder, B., Stantiripop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Tsurganova, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Zhang, L.H., Tsurganova, C., Vogt, J.L., NISC Comparative Sequencing Initiative		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 173068)		
AUTHORS	Green, E.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (15-MAY-2001) NIH Intramural Sequencing Center, 8717		
REFERENCE	3 (bases 1 to 173068)		
AUTHORS	Green, E.D.		


```

TITLE      Direct Submission
JOURNAL    Submitted (25-JUL-2001) NIH Intramural Sequencing Center, 8717
REFERENCE  Grovemont Circle, Gaithersburg, MD 20877, USA
AUTHORS    4 (bases 1 to 173068)
JOURNAL    Green, E.D.
TITLE      Direct Submission
COMMENT    Submitted (10-OCT-2002) NIH Intramural Sequencing Center, 8717
           Grovemont Circle, Gaithersburg, MD 20877, USA
           On Jul 25, 2001 this sequence version replaced gi:14042997.
           ----- Genome Center
           Center: NIH Intramural Sequencing Center
           Center code: NISC
           Web site: http://www.nisc.nih.gov
           Contact: nisc.zoo@nhgri.nih.gov
           ----- Project Information
           Center project name: axa
           Center clone name: 354306

This sequence was finished as follows unless otherwise noted:
all regions were double-stranded, sequenced with an
alternate chemistry, or covered by high quality data
(i.e., phred quality >= 30); an attempt was made to resolve
all sequencing problems, such as compressions and repeats;
all regions were covered by at least one plasmid subclone
or more than one M13 subclone; and the assembly was confirmed
by restriction digest.

CLONE LENGTH: This sequence represents the entire insert of
this clone unless otherwise noted. If there are overlapping
clones, the overlaps are noted in the beginning and end of
the Features section.
           Location/Qualifiers
           source
             1..173068
               /organism="Bos taurus"
               /mol_type="genomic DNA"
               /db_xref="taxon:9913"
               /clone="RP42-354B6"
               /clone_lib="RP42"
               /notes="single clone coverage"
           misc_feature
             145040
               /notes="single clone coverage"

Query Match      59.7%; Score 2131; DB 1; Length 173068;
Best Local Similarity 76.6%; Pred. No. 0;
Matches 2628; Conservative 0; Mismatches 165; Indels 636; Gaps 8;

QY 1  GTGAGGAGGGCGGCGGTGAGGAGAGGTACCTCGTCCAGGTAAGGAGCAGTAGC 60
DB 164074  GTGAGGAGGGCGGCGGTGAGGAGAGGTACCTCGTCCAGGTAAGGAGCAGTAGC 164128
QY 61  TGGCTTTGCTGGAGCAGCGGTAAAGAGATACCCGCGCCAGGTAAGAGAAACCCAG 120
DB 164129  TGGCTTTGCTGGAGCAGCGGTAAAGAGATACCCGCGCCAGGTAAGAGTAACCGAAG 164188
QY 121  TAAGATGTTAGTGTGTTGTGAGAGGCGATCAGAGGGCAGACATCTGAAACCATACACGA 180
DB 164189  TAAGACGGTAGGTGTTGCAAGAGGGCATCAGAGGGCAGACACACTGAAACGATAATCAT 164248
QY 181  GAAAACTAGTCAATCTAATCAGACTAGGACACAGCGCTTGCTTAATCTCAATGAACTAG 240
DB 164249  GAAAACTAGTCAATCTAATCAGACTAGGACACAGCGCTTGCTTAATCTCAATGAACTAG 164308
QY 241  CCATGCCCGTGGGCGAACCAAGATGGCGAGGTCAATGGTGGAGAGATCTGACAGAAATGTG 300
DB 164309  CCATGCCCGTGGGCGAACCAAGATGGCGAGGTCAATGGTGGAGAGTCTGACAGAAATGTG 164368
QY 301  GTCCACTGGAGAGAGGGAAT-GCAAAACCACTTCAGTATCTTTGCTTTGAGAACCCCATGAA 359
DB 164369  GTCCACTGGAGAGGGAATGGCAAAACCACTTCAGTATCTTTGCTTTGAGAACCCCATGAA 164428
QY 360  CAGTATCAAAAGGCAGAAATCATAGGATCTGAAAGAGGAGTCCCGCAGGTCAGTAGTGC 419
DB 164429  CAGTATCAAAAGGCAGAAATCATAGGATCTGAAAGAGGAAATTCGCCAGGTCAGTAGTGC 164488

```

```

QY 420  CCCATATGCTACTGGAGATCAGTGGAGAAATAAATCCAGAAAGAAATGAAGAGATGGAGCC 479
DB 164489  CCAGTAGCTACTGGAGATCAGTGGAGAAATAAATCCAGAAAGAAATGAAGAGATGGAGCC 164548
QY 480  AAGCAAAAAGAAATACCAGCTGCGATGCTGGTGTATATAGCAAGCTCCGATCGTCTG 539
DB 164549  AAGCAAAAAGAAATACCAGCTGCGATGCTGGTGTATATAGCAAGCTCCGATCGTCTG 164608
QY 540  TAAAGAGCAATATTGCAATAGGAACCTGGAATGTGAGTCCATGAATCAAGCAAAATTTGGA 599
DB 164609  TAAAGAGCAATATTGCAATAGGAACCTGGAATGTGAGTCCATGAATCAAGCAAAATTTGGA 164668
QY 600  AGTGGTCAACAAGAGATGGCAAGAGTGAATGTCAACATTTCTAGGAATCAGCAACTAAA 659
DB 164669  AGTGGTCAACAAGAGATGGCAAGAGTGAATGTCAACATTTCTAGGAATCAGCAACTAAA 164728
QY 660  ATGACTGGAAATGGTGAATTTAACTCAGATGACCATTTATATCTTACTTCTGCGGCGAGGA 719
DB 164729  ATGACTGGAAATGGTGAATTTAACTCAGATGACCATTTATATCTTACTTCTGCGGCGAGGA 164788
QY 720  ATCCCTCAGAAAGAAATGGAGTAGCCATCATGGTCAACAAAGAGTCCGAAATGAGTACT 779
DB 164789  ATCCCTTAGAAGAAATGGAGTAGAGATCATGGTCAACAAAGAGTCCGAAATGAGTACT 164848
QY 780  TGGATGCACTCTCAAAAACGACAGATGATCTGTGTTTGTTCACAGGCAAAACCATCAA 839
DB 164849  TGGATGCAATCTCAAAAACGACAGATGATCTGTGTTTGTTCACAGGCAAAACCATCAA 164908
QY 840  TATCACAGTAAATCCAAGTCTATGCCCCCAACAGTAATCTGTAAGAGTGAAGTGAACG 899
DB 164909  TATCACGGTAATCCAAGTCTATGCCCCCAACAGTCAATCAGTAACGCTGAAGTGAATG 164968
QY 900  GTCCTATGAAGACCTACAGACCTTTTAGAATCAACCCCAAAAAGATGTCCTTCTCAT 959
DB 164969  GTTCCATGAAGACCTACAGACCTTTTAGAATCAACCCCAAAAAGATGTCCTTCTCAT 165028
QY 960  TATAGGGGACTGGAATGCAAAAGTAGGAAGCAAAAGAACACCTCGAGAGTAACAGGCAAA 1019
DB 165029  TATAGGGGACTGGAATGCAAAAGTAGGAAGTCAAGAACACCTCGAGAGTAACAGGCAAA 165088
QY 1020  TGGCTTTGGAATACGGAATGAAGCAGGGCAAGAACTAATAGAGTTTTCGCAAGAAATGC 1079
DB 165089  TGGCTTTGGAATACGGAATGAAGCAGGGCAAGAACTAATAGAGTTTTCGCAAGAAAGACGC 165148
QY 1080  ACTGTCATAGCAACACCCCTCTTCCAAACACACAGAGAGAGCTCTACACATGACATC 1139
DB 165149  ACTGTCATAGCAACACCTCTTCCAAACACACAGAGAGAGCTCTACATATGACATC 165208
QY 1140  ACCAGATGGTCAACAACCGAAATCAGATTGATTATATTCTTTGCGCCAAAGATGGAGAG 1199
DB 165209  ACCAGATGGTCAACAACCGAAATCAGATTGATTATATTCTTTGCGCCAAAGATGGAGAG 165268
QY 1200  CTCTATACAGTCAGCAAAAACAGACCGAGAGCTTACTGTGGCTCAGATCATGACTCCT 1259
DB 165269  CTCTATACAGTCAGCAAAAACAGACCGAGAGCTTACTGTGGCTCAGATCATGACTCCT 165328
QY 1260  TATTGCCAAATTCAGACTTAAATTAAGAAAAGTAGGGAAGAACCACTAGATCACTCAGGTA 1319
DB 165329  TATTGCCAAATTCAGACTTAAATTAAGAAAAGTAGGGAAGAACCACTAGATCACTCAGGTA 165388
QY 1320  AGACCTAAATCCAAATCCCTTATGATTATACAGTGGAAAGTGAATATAGATTTAAGGCGCT 1379
DB 165389  TGACCTAAATCCAAATCCCTTATGATTATACAGTGGAAAGTGAATATAGATTTAAGGCGCT 165448
QY 1380  AGATCTGATAGACAGAGTACCTTAATGAATCTAGACAGAGGTTCAATGACATTGTACAGGA 1439
DB 165449  AGATCTGATAGACAGGTTCTGATGAATCTAGACTGAGGTTCAATGATATGTACAGAA 165508
QY 1440  GACAGGGATCGAGAGCCATCCCCATGGAAAAAGAAATGCAAAAAAGCAAAATGCTGCTTGG 1499
DB 165509  GACAGGGATCAAGACCACTCTCTTGGAAAAGAAATGCAAAAAAGCAAAATGCTGCTTGG 165568
QY 1500  GGAGGCGCTTACAAATAGCTGTGAAAAGAGAGAGTGAAGCAAGCAAGGAAAGAAAG 1559

```


AUTHORS

Akhter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Cariaga, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Lalic, P., Lee-Lin, S.-O., Legaspi, R., Maduro, Q.L., Maduro, V.B., Margulies, E.H., Masiello, C., Maskeri, B., McDowell, J., Paguirigan, C., Pearson, R., Portnoy, M.E., Prasad, A., Reddix-bugue, N., Schandler, K., Schueler, M.G., Sison, C., Stantropop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L., Wetherby, K.D., Wiggins, L., Young, A., and Green, E.D.
 NISC Comparative Sequencing Initiative
 Unpublished
 2 (bases 1 to 171153)
 Green, E.D.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 170812 bases at least Q40
 Consensus quality: 170887 bases at least Q30
 Consensus quality: 170927 bases at least Q20
 Insert size: 175000; agarose-fp
 Insert size: 170953; sum-of-contigs
 Quality coverage: 12.55% in Q20 bases; agarose-fp
 Quality coverage: 12.85% in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.

1 88372: contig of 88372 bp in length
 88473 91627: gap of unknown length
 91628 91727: gap of 3155 bp in length
 91728 171153: contig of 79426 bp in length.

FEATURES

source

1. 171153
 /organism="Bos taurus"

/mol_type="genomic DNA"
 /db_xref="taxon:9913"
 /clone="RP42-243A4"
 /clone_lib="RP42"
 1. 88372
 /note="assembly_fragment"
 clone_end:T7
 vector_side:left
 88473. 91627
 /note="assembly_fragment"
 91728. 171153
 /note="assembly_fragment"
 clone_end:SP6
 vector_side:right

Query Match 59.5%; Score 2125.5; DB 1; Length 171153;

Best Local Similarity 77.3%; Pred. No. 0;

Matches 2627; Conservative 0; Mismatches 165; Indels 605; Gaps 10;

QY 1 GTCAGGAGGGGGCAGTGAGGAGAGGTACCTACCTCGTCCAAAGGTAAAGGAGCAGTAGC 60
 DB 35392 GTCAGGAGGGGGCAGTAATGAGGAGA-----TACCCCTCATCCAAGGTAAAGGAGCAGGAC 35446
 QY 61 TCGGCTTTTCTGGAGCAGCCGCTAAAGAGATACCCAGCCCAAGCTAAGAGAAAACCCCAAG 120
 DB 35447 TCGGCTTTTCTGGAGCAGCCGCTGAAGAGATATCCAGCACAAGGTAAAGAGAAAACCCCAAG 35506
 QY 121 TAAGATGTAGTGTGTGAGAGGGCATCAGAGGGCAGACATCTGTAACCATACATACGCA 180
 DB 35507 TAAGACGTAGTGTGTGCAAGAGGGCATCAGAGGGCAGACAACTGAAACCATATCTACA 35566
 QY 181 GAATACTAGTCAATCTAATCACACTAGGACCAAGCCCTTGTAACTCAATCAATGAACTAAG 240
 DB 35567 TAAACTAGTCAATCTAATCACACTAGGACCAAGCCCTTGTCTAACTCAATGAACTAGG 35626
 QY 241 CCAATCCGCTGGGGCAACCCAGATGGCAGGTGATGGTGGAGAGATCTGACAGATGTG 300
 DB 35627 CCAATCCGCTGGGGCAACCCAGATGGCAGGTGATGGTGGAGAGATCTGACAGATGTG 35686
 QY 301 GTCACACTGAGAAGGGGAAT-GCAAAACCACTTCAGTATTCTTTCCTTGAGAACCCCATGAA 359
 DB 35687 GTCACACTGAGAAGGGGAATGCAAAACCACTTCAGTATTCTTTCCTTGAGAACCCCATGAA 35746
 QY 360 CAGTATGAAAAGGCAAAAATGATAGGATATCTGAAAGAGAACTCCCAAGGTCAAGTGTGC 419
 DB 35747 CAGTATGAAAAGGCAAAAATGATAGGATATCTGAAAGAGAACTCCCAAGGTCAAGTGTGC 35805
 QY 420 CCATATCTCTACTGGAGATCTAGTGGAGAAATACTCCAGAGAGAAATGAGAGATGGAGCC 479
 DB 35806 CCAATATCTCTACTGGATATCAGTAGAGAAATACTCCAGAGAGAAATGAGAGATGGAGCC 35865
 QY 480 AAAGCAAAAAGAAATACCCAGCTGTGGATGTGACTGGTGATATAAGCAAGGTCCGATGCTG 539
 DB 35866 AAAGCAAAAAGAAATACCTAGTGTGGATGTGACTGGTGATATAAGCAAGGTCCGATGCTG 35925
 QY 540 TAAAGCAATATTGCAATAGGAACCTGGAATGTGAGTCCATGATCAAGCAAAATTGGA 599
 DB 35926 TAAAGCAATATTGCAATAGGAACCTGGAATGTGAGTCCATGATCAAGCAAAATTGGA 35985
 QY 600 AGTGTGCAAAACAAGAGATGGCAAGAGTGAATGTCAACATTCTAGCAATCAGCAACTAAA 659
 DB 35986 AGTGTGCAAAACAAGAGATGGCAAGAGTGAATGTCAACATTCTAGCAATCAGCAACTAAA 36045
 QY 660 ATGGACTGGAATGGGTGAATTTAACTCAGATGACCATTTATCTACTACTCTCGGGCAGGA 719
 DB 36046 ATGGACTGGAATGGGTGAATTTAACTCAGATGACCATTTATCTACTACTCTCGGGCAGGA 36105
 QY 720 ATCCCTCAGAGAAATGGAGTAGCCATCATGGTCAAGAAAGATCCCAAAATGGAGTACT 779
 DB 36106 ATCCCTCAGAGAAATGGAGTAGCCATCATGGTCAAGAAAGATCCCAAAATGGAGTACT 36165
 QY 780 TGGATGCAAGTCTCAAAAAACGACAGAATGATCTGTGTTTTCCTCAAGGCAAAACCATTCAA 839

Db	36166	TGGATGCAATCTCAAGATGACAGATGATCTCTGTTGTTTCCAAAGGCAAAACCATTCAG	36225
Qy	840	TATCACAGTAAATCCAACTCTATGCCCCCAACAGTAAATGCTGAAGAAAGCTGAACTTGAACG	899
Db	36226	TATCACAGTAAATCCAACTCTATGCCCCCAACAGTAAATGCTGAAGAAAGCTGAACTTGAACG	36285
Qy	900	GTCCTATGAGACCTCAAGACCTTTTAGAACTTAACACCCCAAAAAGATGCTCTCTCAT	959
Db	36286	GTTCTATCAAGACCTCAAGACCTTTTAGAACTTAACACCCCAAAAAGATGCTCTCTCAT	36345
Qy	960	TATAGGGAGCTGGAATCCAAAGTAGAAGCAAGAAACACCTGGAGTACACGGCAAAAT	1019
Db	36346	TATAGGGAGCTGGAATCCAAAGTAGAAGTCAAGAAACACCTGGAGTACACGGCAAAAT	36405
Qy	1020	TGGCCTTTGGAAATACGGAATGAAGCAGGGCAAGACTAATAGATTTTGGCCAGAAAATGC	1079
Db	36406	TGACCTTTGGAAATACGGAATGAAGCAGGGCAAGACTAATAGATTTTGGCCAGAAAATGC	36465
Qy	1080	ACTGGTCATAGCAACACCTCTTCCAAACACACAGAGAGACTCTACACATGGACATC	1139
Db	36466	ACTGGTCATAGCAACACCTCTTCCAAACACACAGAGAGACTCTACACATGGACATC	36525
Qy	1140	ACCAGATGGTCAACACCGAAATCAGATTGATTATATTTCTTTGGCAGCCAAAGATGGAGAAG	1199
Db	36526	AGCAGATGGTCAACACCGAAATCAGATTGATTATATTTCTTTGGCAGCCAAAGATGGAGAAG	36585
Qy	1200	CTCTATACAGTACGCAAAAACAGACAGAGGCTTACTGTGGCTCAGATCATGAATCCT	1259
Db	36586	CTCTATACAGTACGCAAAAACAGACAGAGGCTTACTGTGGCTCAGATCATGAATCCT	36645
Qy	1260	TATTGCCAAATTCAGACTTAAATTTGAAGAAAGTAGGAAACCACTAGATCACTCAGGTA	1319
Db	36646	TATTGCC- AATTAGACTGAAATTTGAAGAAAGCAGGGAAACCACTAGACCATTGAGTA	36704
Qy	1320	AGACCTAAATCCAAATCCCTTATGATTATACAGTGGAAAGTGAAGAAATAGATTTAAGGGCT	1379
Db	36705	TGACCTAAATCCAAATCCCTGATGATTATACAGTGGAAAGTGAAGAAATAGATTTAAGGGCT	36764
Qy	1380	AGATCTGATAGACAGAGTACTTAATGAATATGACAGAGGTTTCATGACATTGTACAGGA	1439
Db	36765	AGATCTGATAGACAGAGTACTTAATGAATATGACAGAGGTTTCATGACATTGTACAGGA	36824
Qy	1440	GACAGGATCGAGACCTCCCAATGGAAAGAAATGCCAAAGAAATGGCTGTCTCG	1499
Db	36825	GACAGGATCAAGACCATCCCAATGGAAAGAAATGCCAAAGAAATGGCTGTCTCGA	36884
Qy	1500	GGAGGCTTACAAATAGCTGTGAAGAGAGAGTGAAGAGCAAGGAAAGAAAGGAAG	1559
Db	36885	GGAGGCTTACAAATAGCTGTGAAGAGAGAGTGAAGAGCAAGGAAAGAAAGGAAG	36944
Qy	1560	ATAAAGCATCTGAATCAGAGTTCCAAA	1588
Db	36945	ATAAAGCATCTGAATCAGAGTTCCAAA	37004
Qy	1589	-----	1588
Db	37005	CAGCGATCAATGCAAGAAATAGAGGAAACAAACAGAAATGGGAAAGTAGGGATCTCTTCA	37064
Qy	1589	-----	1588
Db	37065	AGAAATCAGAGATACCAAGGAACATTTTCATGCAAGATGAGCTCGATAAAGGACAGAA	37124
Qy	1589	-----	1588
Db	37125	ATGGTATGAATCAACAGAGCAGAGATATTAAGAAAGAGATGGCAAGAAATACACAGAGA	37184
Qy	1589	-----	1588
Db	37185	ACTGTACAAAAAGACCTTCGTGACCCAGATAATCATGATAGTGTGATCACTGACCTAGA	37244
Qy	1589	-----	1588
Db	37245	GCCAGACATCTGGAATGTGAAGTCAAGTGGGCTTTAGAAAGCATCACTATGAACAAAGC	37304
Qy	1589	-----	1588
Db	37305	TAGTGGAGTGAATTCAGATTCCAGTTGAGCTGTTGCAAGTCTGAAAGATAATGCTGGAA	37364
Qy	1589	-----	1588
Db	37365	AGTGTGCACTCAATATATCCAGCAAAATTTGAAAACTCAGCAGTGGCCACAGGACTGGAA	37424
Qy	1589	-----	1588
Db	37425	AGGTTCAGTTTTCATTTCCAAATCCCAAGAAAGCAATGCTCCAAACTACCG	37484
Qy	1589	-----	1588
Db	37485	CACAATTTGCACTCATCTCACACGCTAGTAAATGATGCTCAAAATTTCTCAAGCCAGGCT	37544
Qy	1589	-----	1588
Db	37545	TCAGCAATATCGCAACCGTGAACCTTCTGATATTTCAAGCTGTTTGAAGAAAGCAGAGG	37604
Qy	1630	AACCAAGACCAAAATTTGCCAACATCTCTGATCATGGAAGAAAGCAAGAGTTTCCAGAA	1689
Db	37605	AACCAAGACCAAAATTTGCCAACATCTCTGATCATGGAAGAAAGCAAGAGTTTCCAGAA	37664
Qy	1690	AAACATCTATTTCTGCTTTTATTGACTATGCAAAAGCCTTTGACTGTGGGTCAACAATA	1749
Db	37665	AAACATCTATTTCTGCTTTTATTGACTATGCAAAAGCCTTTGACTGTGGGTCAACAATA	37724
Qy	1750	ACTGTGGAATAATCTGAAAGGATGGGAATACCAAGACCACTGACCTGACTCTTTGAAAA	1809
Db	37725	ACCGTGGAAATTTCTGAAAGAGATGGGAATACCAAGACCGCTTGATCTGCTTTGAGAA	37784
Qy	1810	TTTGTATGCAAGTCAAGAAAGCAAGTTAGAACTGGAACATGGAACCAACAGACTGGTTCCA	1869
Db	37785	TTTGTATGCAAGTCAAGAAAGCAAGTTAGAACTGGAACATGGAACCAACAGACTGGTTCCA	37844
Qy	1870	AGTAGGAAAGGAGTATGTCAAAGCTGTATATTGTCACCCGGCTGTTTAACTTCTATGC	1929
Db	37845	AGTAGGAAAGGAGTATGTCAAAGCTGTATATTGTCACCCGGCTGTTTAACTTCTATGC	37904
Qy	1930	AGAG-ACATCATGAGAAACCGCTGGCTGGAAAGAGCAAGCTGGAATCAAGATTGCCGG	1988
Db	37905	AGAGTACATCATGAGAAACCACTGGACTGGAATCAAGAAAGCTGGAATCAAGATTGCCGG	37964
Qy	1989	GAGAAATAGCAATAACCTCAGATATGAGATGATACCAACCCCTTATGCAAGAAAGTGAAGA	2048
Db	37965	GAGAAATAGCAATAACCTCAGATATGAGATGATACCAACCCCTTATGCAAGAAAGTGAAGA	38024
Qy	2049	GGAACTAAAAAGCCTCTTGATGAAGGTGAAGAGAGAGTGAAGAAAGTTGGCTTAAAGCT	2108
Db	38025	GGAACTAAAAAGCCTCTTGATGAAGGTGAAGAGTGAAGAAAGTTGGCTTAAAGCT	38084
Qy	2109	CAACATTCAGAAACAGAGATCATGGCATCTGGTCCCTCATCTCATGGAAGATAGATGG	2168
Db	38085	CAACATTCAGAAACAGAGATCATGGCATCTGGTCCCTCATCTCATGGAAGATAGATGG	38144
Qy	2169	GGAAACAGTGAAGACAGTGTGAGACTTTTATTTTGGGGGCTCCAAAATCACTGCAAGT	2228
Db	38145	GGAAACAGTGAAGACAGTGTGAGACTTTTATTTTGGGGGCTCCAAAATCACTGCAAGT	38202
Qy	2229	GGTGAATGACCCATGAAATTAAGACACTTACTCTTCTGGAAGAAAGTTA-ACCAAC	2286
Db	38203	GGTGAATGACCCATGAAATTAAGACACTTACTCTTCTGGAAGAAAGTTAAGCAAC	38262
Qy	2287	CTAGATACATATTTGAAAGCAGAGACATTACCTTGGCAACAAAGCCCATCTAGTCAAG	2346
Db	38263	CTAGATACATATTTGAAAGCAGAGACATTACTTTGCCAACAAGGTCCTCTAGTCAAG	38322
Qy	2347	GCTATGTTTTTTCAGTGGTTCATGTATGGATGTGAGAGTTGGAGTGTGAAGAAAGCTGAG	2406
Db	38323	GCTATGTTTTTTCAGTGGTTCATGTAT-GATGTGAGAGTTGGAGTGTGAAGAAAGCTGAG	38381

QY	2407	CACTGAAGAATTGATGCTTTTGAAC	TGTGGTCTTGGAGAAGACTCTT	GAGAGTCCCTTGG	2466
Db	38382	CGCCGAAGAATTGATGCTTTTGAAC	TGTGGTCTTGGAGAAGACTCTT	GAGAGTCCCTTGG	38441
QY	2467	ACTGCAAGGAGATCCAAACAGATCCAT	TTCTGAAGGAGATCAGCCCTGGGAT	TTCTTTGGAA	2526
Db	38442	ACTGCAAGGAGAT-CAACCAATCCAT	TTCTGAAGGAGATCAGCCCTGGGAT	TTCTTTGGAA	38500
QY	2527	GGAACTGATCTAAAGCTGAAACTCCAG	TACTTTGGCCACCTCATCAGAAGAGCT	GACTCA	2586
Db	38501	GGAACTGATCTAAAGGTGAACTCCAG	TACTTTGGCCACCTCATCAGGAAGAGT	TGACTCA	38560
QY	2587	CTGGAAAGACCCCTGATGCTGGAGGGAT	TGGGGGACAGAGAGAGGGGACGACAGAGG	2646	
Db	38561	TTGGAAAGACTCTGATGCTGGAGGGAT	TGGGGGACAGAGAGAGGGGACGACAGAGG	38620	
QY	2647	ATCAGATGGCTGATGGCATCACTGACTCGAT	CGACGTCAGTCTGGGTCACTCCCTGGAG	2706	
Db	38621	ATCAGATGGCTGATGGCATCACTGACTCGAT	CGACGTCAGTCTGGGTCACTCCGGAG	38680	
QY	2707	TTGGTGATGACAGGGAGGCGCTCTCCTCG	GGGCATTCATGGGGTCAAAAGAGTTGGACA	2765	
Db	38681	TTGGTGATGACAGGGCGCGCTCTCCTCG	GGGCATTCATGGGTCCCAAAGATCGGACA	38740	
QY	2767	CGACTGAGCAACTGAACTGAACTGATACTG	AAAA	2803	
Db	38741	CGACTGAGTGACTGATCTGATCTGATCTG	GTCTTAAA	38777	

RESULT 8					
AC142564					
LOCUS	AC142564	145088 bp	DNA	linear	HTG 02-MAY-2003
DEFINITION	Bos taurus clone RP42-56B14, WORKING DRAFT SEQUENCE, 2 ordered pieces.				

ACCESSION AC142564
VERSION AC142564.2 GI:30315712
KEYWORDS HTG: HTGS PHASE2: HTGS DRAFT.

SOURCE	ORGANISM
Bos taurus (cow)	Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;
Bovidae; Bovinae; Bos.

REFERENCE 1 (basses 1 to 145088)

REFERENCE
AUTHORS

1. (bases 1 to 145088)

Atanocellis, A., Avela, K., Beckstrom-Sternberg, S. M., Benjamin, B.,
Blakesley, R. W., Bouffard, G. G., Brinkley, C., Brooks, S., Cariga, K.,
C. G., Coleman, B., Coleman, H., Engle, J., Granice, S., Guan, X.,
Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Hu, P.,
Hurdle, B., Idol, J. R., Karlins, E., Kwong, P., Lari, P., Lee-Lin, S.-O.,
Legaspi, R., Maduro, Q. L., Maduro, V. B., Magulies, E. H., Mastello, C.,
Masker, B., McDowell, J., Paquirigan, C., Pearson, R., Portney, M. E.,
Prasad, A., Reddix-Bugue, N., Schandler, K., Schueler, M. G., Shah, K.,
Sison, C., Stantripop, S., Thomas, P. J., Tsipouri, V.,
Voor, J. L., Wertheby, K. D., Wiggins, L., Young, A. and Green, E. D.

TITLE	NISC Comparative Sequencing Initiative
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 145088)

2 (boxed 1 to 110000)
Green, E.D.
Direct Submission
Submitted (04-Apr-2003) NIH Intramural Sequencing Center, 8717
JOURNAL
Groveont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 145088)
REFERENCE

REFERENCE	3 (bases 1 to 143068)
AUTHORS	Green, E.D.
TITLE	Direct Submission
JOURNAL	Submitted (02-MAY-2003) NIH Intramural Sequencing Center, 8717 Government Circle, Gaithersburg, MD 20877, USA
COMMENT	On May 2 2003 this sequence version replaced id:29540599.

COMMENT
ON May 2, 2003 this sequence version replaced gr:29340333

Genome Center

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: <http://www.nisc.nih.gov>
Contact: nisc_zoo@nih.gov

----- Project Information

Center project name: djr

Center clone name: 056B14

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 144975 bases at least Q40
Consensus quality: 144984 bases at least Q30
Consensus quality: 144986 bases at least Q20
Insert size: 159000; agarose-fp
Insert size: 14988; sum-of-contigs
Quality coverage: 14.02x in Q20 bases; agarose-fp
Quality coverage: 15.37x in Q20 bases; sum-of-contigs

- * NOTE: This is a 'working draft' sequence. It currently
- * consists of 2 conigs. Gaps between the conigs
- * are represented as runs of N. The order of the pieces
- * is believed to be correct as given, however the sizes
- * of the gaps between them are based on estimates that have
- * provided by the submittor.

provided by the Government.

- * This sequence will be replaced
- * by the finished sequence as soon as it is available and
- * the accession number will be preserved.

Accession number	Gene	Size (bp)	Accession number	Gene	Size (bp)
U00096	32951	32951	U00096	33051	33051
U00096	32952	32952	U00096	33052	33052
U00096	32953	32953	U00096	33053	33053
U00096	32954	32954	U00096	33054	33054
U00096	32955	32955	U00096	33055	33055
U00096	32956	32956	U00096	33056	33056
U00096	32957	32957	U00096	33057	33057
U00096	32958	32958	U00096	33058	33058
U00096	32959	32959	U00096	33059	33059
U00096	32960	32960	U00096	33060	33060
U00096	32961	32961	U00096	33061	33061
U00096	32962	32962	U00096	33062	33062
U00096	32963	32963	U00096	33063	33063
U00096	32964	32964	U00096	33064	33064
U00096	32965	32965	U00096	33065	33065
U00096	32966	32966	U00096	33066	33066
U00096	32967	32967	U00096	33067	33067
U00096	32968	32968	U00096	33068	33068
U00096	32969	32969	U00096	33069	33069
U00096	32970	32970	U00096	33070	33070
U00096	32971	32971	U00096	33071	33071
U00096	32972	32972	U00096	33072	33072
U00096	32973	32973	U00096	33073	33073
U00096	32974	32974	U00096	33074	33074
U00096	32975	32975	U00096	33075	33075
U00096	32976	32976	U00096	33076	33076
U00096	32977	32977	U00096	33077	33077
U00096	32978	32978	U00096	33078	33078
U00096	32979	32979	U00096	33079	33079
U00096	32980	32980	U00096	33080	33080
U00096	32981	32981	U00096	33081	33081
U00096	32982	32982	U00096	33082	33082
U00096	32983	32983	U00096	33083	33083
U00096	32984	32984	U00096	33084	33084
U00096	32985	32985	U00096	33085	33085
U00096	32986	32986	U00096	33086	33086
U00096	32987	32987	U00096	33087	33087
U00096	32988	32988	U00096	33088	33088
U00096	32989	32989	U00096	33089	33089
U00096	32990	32990	U00096	33090	33090
U00096	32991	32991	U00096	33091	33091
U00096	32992	32992	U00096	33092	33092
U00096	32993	32993	U00096	33093	33093
U00096	32994	32994	U00096	33094	33094
U00096	32995	32995	U00096	33095	33095
U00096	32996	32996	U00096	33096	33096
U00096	32997	32997	U00096	33097	33097
U00096	32998	32998	U00096	33098	33098
U00096	32999	32999	U00096	33099	33099
U00096	33000	33000	U00096	33100	33100
U00096	33001	33001	U00096	33101	33101
U00096	33002	33002	U00096	33102	33102
U00096	33003	33003	U00096	33103	33103
U00096	33004	33004	U00096		

FEATURES

source

```

/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="RP42-56B14"
/clone_lib="RP42"

```

misc feature

```
misc_feature
T: 0.5251
/note="assembly_fragment
clone_end:T7
vector_side:left"
```

misc feature

```
misc_reacure
0002: 00000000
/note="assembly_fragment
clone_end:SP6
vector_side:right"
```

misc feature

```
misc_feature
      .clone_overlap()
    /note="clone overlaps with GenBank Accession Number  
AC141853 clone RP42-374H3 (center project name dip)"
```

```
Query Match      59.4%; Score 2123.49; DB 1; Length 145088;
Best Local Similarity 75.1%; Pred. No. 0;
Matches 2699; Conservative 0; Mismatches 280; Indels 615; Gaps 10;
```

1 GTCAGGAAGGGCGGCAGTGAGGAGAGGTACCTACCCCTCGTCCAAGGTAGGAGCAGTAGC 60

[illegible]

80704 GTCAGGTAAGGGCCCGCTGGTGGACAGA-----TACCCTCCTCATCCATCTGTGTTTCTGCAGCTATTGCG
DD

61 TCGCTTGCTGCAAGCACCGCTAAGSAGATACCCCACGCCCAAGGTAAAGAGAACCCCCAAG 120
CV

QY TGGCTTTGCTGGAGCAGCCGTAATAGAGAATACTCCACGCCCAAGGTAAAGAGAGTCCCATG
69

DB 80759 TGTGCTTTGCTGGAGCAGCTGTGAAGAGATACCCACGCCCCAAGGTAAGAGAGGACCCCAAG 90019

121 TAA GAT GGT AGT GTT GTG AGA GGG CAT CAG AGG CAG ACA TACT GAA ACC ATA CAC GCA 180

83030	DB	TAACCTGTGAAAAAATTCTGAAGAAGATGGGAATACAGACCACTGATCTGCCTCTTTGAG	83089
1807	QY	AAATTTGTATGCAGGTTCAGGAAGCAACAGTTTAGAACTGGACATGGAAACAACAGACTGGTT	1866
83090	DB	AAATTTGTATGCAGGTTCAGGAAGCAACAGTTTAGAACTGGACATGGAAACAACAGACTGGTT	83149
1867	QY	CCAAGTAGAAAAGAGTATGTCAAGCGCTGTATATTGTCAACCGGCTGTGTTAACTTCTA	1926
83150	DB	CCTAATAGAAAAGAGTACGTCAAGGCTGTATACTGTAAACCTGTTTATTTTGCCTTATA	83209
1927	QY	TGCAGAG-ACATCATGAGAAACGCTTGGCTTGGAAAGAACACAAAGCTGGAATCAAGATTGC	1985
83210	DB	TTTCAAGTACATCATGAGAAACGCTGGACTGGAGAGGACACAAGCTGGAATCAAGATTGC	83269
1986	QY	CGGGAGAAATAGCAATAACCTCAGATATGCGATGATACCAACCTTATGGCAGAAAGTGA	2045
83270	DB	CGGGAGAAATATCAATAACCTCAGATATGCGATGATACCAACCTTATGGCAGAAAGTGA	83329
2046	QY	AGAGGAACATAAAGACCTCTTCATCAAGGTGAAGAGGAGAGTGAAGAAAGTTGGCTTAAA	2105
83330	DB	AGAGGAACTATAAAGACCTCTTCATCAAGGTGAAGAGGAGAGTGAAGAAAGTTGGCTTAAA	83389
2106	QY	GCTCAACATTTCAAAAAAGAAATCATGGCATCTGGTCCCATCTTCATGGGAAATAGA	2165
83390	DB	GCCCAACATTCAGAAAAAGAAATCATGGCATCCGGTCCCATCTTCATGGGAAATAGA	83449
2166	QY	TGGGGAAAAGTGGAAAACAGTCTCAGACTTTATTTTTTTGGGGGGCTCCAAATCACTGCA	2225
83450	DB	TGGGGAAAAGTGGAAAACAGTCTCAGACTTTATTTTTTTCTGGGCTCCAAATCACTGCA	83507
2226	QY	GATGGTGACTGCGAGCCATCAAAATTTAAAAGACACTTACTCTTTGGAAGAAAAGTTA--ACC	2283
83508	DB	GATGGTGACTGCGAGCCATCAAAATTTAAAAGATGCTTGTCTCTTGGAAAGAAAAGTTATGACC	83567
2284	QY	AACCTAGATAGCAATTTGAAAAGCAGAGACATTAAGCTTGCCTTCCCAACAAAGCCCACTAGTC	2343
83568	DB	AACCTAGATAGCAATTTCAAAAAGCAGAGACATTAAGCTTGCCTTCCCAACAAAGCTCTAGTC	83627
2344	QY	AAGGCTATGGTTTTTCCAGTGGTCACTGATGTGATGTGAGAGTTGAGCTGTCAAGAAAGCT	2403
83628	DB	AAGGCTATGGTTTTTCCAGTGGTCACTGATGTGATGTGAGAGTTGGGCTGTCAAGAGGCT	83687
2404	QY	GAGCATGAAGAAATGATGCTTTTTGAACTGTGTGGTGTGGAGAAAGCTCTTGAGAGTCCCT	2463
83688	DB	GAGCG-CCGAAGAAATGATGCTTTTTGAACTGTGTGGTGTGGAGAAAGCTCTTGAGAGTCCCT	83747
2464	QY	TGGAC-TGCAAGAGAGATCAACACAGTCCATCTCTGAAGGAGATCAAGCTGGATTTCTTT	2522
83748	DB	TGACATTGCAAGAGAGATCAACACAGTCCATCTCTGAAGGAGATCAAGCTGGATTTCTTT	83807
2523	QY	GGAGGAATGATGCTAAAGCTGAACTCCAGTACTTTGGCCACCTCATGCAAGAGCTGA	2582
83808	DB	GGAGGAATGATGCTAAAGCTGAACTCCAGTACTTTGGCCACCTCATGCAAGAGCTGA	83867
2583	QY	CTCACTGGAAAAAGACCTGATGCTGGAGGGATTTGGGCGAGGAGAGAGGGACGACA	2642
83868	DB	C-CCTTGGAAAAAGACTCTGATCTCTGGAGGGATTTGGGGCAGAGAGAGAGGGACGACA	83926
2643	QY	GAGGATCAGATGGCTGGATGGCATCACTGACTCGATGGACGTTGAGTCTGGGTGAACCTCCT	2702
83927	DB	GAGGATCAGATGGCTGGATGGCATCACTGACTCGATGGACGTTGAGTCTGGGTGAACCTCCT	83986
2703	QY	GGAGTTGGTGAATGGAACAGGAGCGCTGTCTCTGGGGGATTCATGGGGTCAAGAGAGTTG	2762
83987	DB	GTAGTTGGTGAATGGAACAGGAGCGCTGTCTCTGGGGGATTCATGGGGTCAAGAGAGTTG	84046
2763	QY	GACACGACTCAGCAACTGAACTGAACTGTAAGTCTTACTGAAACCTTAGTAGTTTATATTAC	2822
84047	DB	GACACGACTCAGTCACTGATCTGATCTGATCTGATCTGATCACTTTTGGTCATTTTATAA	84106
2823	QY	TCGAAAAATAGTAATTTTATATGATTTCAAAATTTTTCATTAATGTTGGTTAAGATAATA	2882
84107	DB	GACTTTTATATATATATGATTTTATTTTCAAACTCTTATGGGATAGATTGTT	84166

Qy	2883	AGATTTTCAAATGGATTTTTTATCTTTGATTTTTCTCTACTCTATTTAAATTTTGGGATTTTA	2942
Db <td>84167 <td>AATACCTTTTACAGTTGAGGAAATTGATGCCAAAGCTGAATGACTTTCCAGGGTCT <td>84226</td> </td></td>	84167 <td>AATACCTTTTACAGTTGAGGAAATTGATGCCAAAGCTGAATGACTTTCCAGGGTCT <td>84226</td> </td>	AATACCTTTTACAGTTGAGGAAATTGATGCCAAAGCTGAATGACTTTCCAGGGTCT <td>84226</td>	84226
Qy <td>2943 <td>ACTATTTCTCAATGACTTGTATTTCTTAATATTTTACTTATCTATTTACTTTTACITTTA <td>2996</td> </td></td>	2943 <td>ACTATTTCTCAATGACTTGTATTTCTTAATATTTTACTTATCTATTTACTTTTACITTTA <td>2996</td> </td>	ACTATTTCTCAATGACTTGTATTTCTTAATATTTTACTTATCTATTTACTTTTACITTTA <td>2996</td>	2996
Db <td>84227 <td>AATTACTAATAAATCTGCGAGAACCTAAGATATATATAAGGTATCTGGATTCA <td>84280</td> </td></td>	84227 <td>AATTACTAATAAATCTGCGAGAACCTAAGATATATATAAGGTATCTGGATTCA <td>84280</td> </td>	AATTACTAATAAATCTGCGAGAACCTAAGATATATATAAGGTATCTGGATTCA <td>84280</td>	84280

RESULT 9
AC141853
LOCUS AC141853 192431 bp DNA linear HTG 12-APR-2003
DEFINITION Ros taurinus clone RP42-374H3. WORKING DRAFT SEQUENCE, 3 ordered

RESULT 9	AC141853	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
----------	----------	-------	------------	-----------	---------	----------	--------	----------

AC141853 192431 bp DNA linear HTG 12-APR-2003
Bos taurus clone RP42-374H3, WORKING DRAFT SEQUENCE, 3 ordered
pieces.
AC141853
AC141853.2 GI:29824448
HTG: HTGS PHASE2: HTGS DRAFT.

SOURCE	ORGANISM	Bos taurus	Bos taurus (cow)
	Bos taurus		
	Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;		
	Bovidae; Bovinae; Bos.		

REFERENCE

AUTHORS

1. (bases 1 to 192431)
Evinac, J., Duvina, J.,
Antonsellis, A., Ayelle, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G., Brinkley, C., Brooks, S., Cariaga, K.,
Chu, G., Coleman, B., Coleman, H., Engle, J., Grant, S., Guan, A.,
Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Hu, P.,
Hurlb, B., Idol, J.R., Karlins, E., Kwong, P., Laric, P., Lee-Lin, S.-Q.,
Legaspi, R., Maduro, Q.L., Maduro, V.B., Margulies, E.H., Masello, C.,
Maski, B., McDowell, J., Pagnirican, C., Pearson, R., Portnoy, M.E.,
Prasad, A., reddix-Dugue, N., Pagnirican, C., Schuele, M.G., Shah, K.,
Sison, C., Stantripop, S., Thomas, J.W., Thomas, P.J., Taipouri, V.,
Yee, T., Weatherly, K.D., Wiggins, I., Young, A. and Green, E.D.

TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
-------	---------	-----------	---------	-------	---------	-----------	---------	-------	---------	---------

2 (bases 1 to 192431)
Green, E.D.
Direct Submission
Submitted (19-MAR-2003) NIH Intramural Sequencing Center, 8717
Grovmont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 192431)

Direct Submission
Submitted (12-APR-2003) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
On Apr 12 2003 this sequence version replaced gi:29124103.

COMMENT

----- Genome center. -----
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: <http://www.nisc.nih.gov>
 Contact: nisc.zoo@nhgri.nih.gov
 ----- Project information -----
 Center project name: djp
 Center clone name: 374H03

The sequence data in this record represents an 'enhanced' version of a phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been revised and the contigs have been reassembled. The contigs were established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

```

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 192108 bases at least Q40

```


AC108887
 AC108887.2 GI:22262483
 HTG: HTGS_PHASE2; HTGS_DRAFT.
 SOURCE
 Bos taurus
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos
 1 (bases 1 to 173207)
 Akter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
 Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C.,
 Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J.,
 Haghighi,P., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E., Laric,P.,
 Lee-Lin,S.-Q., Legaepi,R., Maduro,Q.L., Maduro,V.B.,
 Margulies,E.H., Mastello,C., Maskeri,B., Mastrian,S.D.,
 McCloskey,J.C., McDowell,J., Paguirigan,C., Pearson,R.,
 Portnoy,M.E., Prasad,A., Schueler,M.G., Stantripop,S., Thomas,J.W.,
 Thomas,P.J., Touchman,J.W., Taurgeon,C., Vogt,J.L., Walker,M.A.,
 Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.
 NISC Comparative Sequencing Initiative
 Unpublished
 2 (bases 1 to 173207)
 Green,E.D.
 Direct Submission
 Submitted (01-FEB-2002) NIH Intramural Sequencing Center, 8717
 Government Circle, Gaithersburg, MD 20877, USA
 3 (bases 1 to 173207)
 Green,E.D.
 Direct Submission
 Submitted (15-AUG-2002) NIH Intramural Sequencing Center, 8717
 Government Circle, Gaithersburg, MD 20877, USA
 On Aug 15, 2002 this sequence version replaced gi:18464063.
 ----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: <http://www.nisc.nih.gov>
 Contact: nisc_zoo@nigr.nih.gov
 ----- Project Information
 Center project name: cba
 Center clone name: 014F06

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics
 Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 172823 bases at least Q40
 Consensus quality: 172849 bases at least Q30
 Consensus quality: 172900 bases at least Q20
 Insert size: 164000; agarose-fp
 Insert size: 172907; sum-of-contigs
 Quality coverage: 13.19x in Q20 bases; agarose-fp
 Quality coverage: 13.51x in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and

* the accession number will be preserved.
 * 1 42204: contig of 42204 bp in length
 * 42205 42304: gap of unknown length
 * 42305 46114: contig of 4110 bp in length
 * 46115 46514: gap of unknown length
 * 46515 109660: contig of 63146 bp in length
 * 109661 109760: gap of unknown length
 * 109761 173207: contig of 63447 bp in length.
 Location/Qualifiers
 1..173207
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9913"
 /clone="RP42-14F6"
 /clone_lib="RP42"
 1..42204
 /note="assembly_fragment"
 clone_end:77
 vector_side:left
 1..24229
 /note="clone overlaps with GenBank Accession Number AC108894 clone RP42-571C2 (center project name axz)"
 42305..46414
 /note="assembly_fragment"
 46515..109660
 /note="assembly_fragment"
 109761..173207
 /note="assembly_fragment"
 clone_end:SP6
 vector_side:right
 145214..173207
 /note="clone overlaps with GenBank Accession Number AC108891 clone RP42-252K20 (center project name cbb)"
 Query Match 59.0%; Score 2108.1; DB 1; Length 173207;
 Best Local Similarity 77.1%; Pred. No. 0;
 Matches 2613; Conservative 0; Mismatches 164; Indels 613; Gaps 10;
 QY 1 GTCAGAGAGGCGGCGAGTGGAGAGGTTACCTCCCTCGTCCCAAGGTAAGGAGCAGTAGC 60
 Db 30890 GTCAGAGAGGCGGCGGCTGAGGAGA-----TAACCTCATCCAGGTAAGGAGCAATGGC 30944
 QY 61 TCGCGTTCTTGGAGACGCCCTTAAGAGATACCCACGCCCAAGGTAAGAGAAACCCCAAG 120
 Db 30945 TGTGTTTCTGTCAGCAGCCGCTGAAGAGATACCCACACCCCAAGTAAAGAGAAACCCCAAC 31004
 QY 121 TAAGATGTTAGTGTGTGTGAGAGGCGCATCAGAGGCGAGACATCTGAAACCATACACGCA 180
 Db 31005 TAAGATGTTAGTGTGTGCAAGAGGCGCATCAGAGGCGAGACATCTGAAACCATACTACA 31064
 QY 181 GAAACTAGTCAATCTAATCACTACCTAGGACACACAGCCCTTGTCTAACTCAATGAACTAAG 240
 Db 31065 GAAACTAGTCAATCTAATCACTACCTAGGACACACAGCCCTTGTCTAACTCAATGAACTAAG 31124
 QY 241 CCATGC-CCGTGGGGCCACCCCAAGATGGGAGGTCATGGTGGAGAGATCTGCAGAGATGT 299
 Db 31125 CCATGCTGTGTGGGGCAACCCCAAGATGGGCGGGTCATGGTGGAGAGATCTGCAGAGATGT 31184
 QY 300 GGTCCACTGGAGAGAGGGAAT-GCAAAACCACTTCACTAGTATTCTTGCCTTGAGAACCCCATGA 358
 Db 31185 AGTCCACTGGAGAGGGAATGGGCAAAACCACTTCACTAGTATTCTTGCCTTGAGAACCCCATGA 31244
 QY 359 ACAGTATGAAAGGCAAAATGATAGGATCTGAAAGAGGAACTCCCGAGTCTAGTAGTG 418
 Db 31245 ACAGTATGAAAGGCAAAATGATAGGATCTGAAAGAGGAACTCCCGAGTCTAGTAGTG 31304
 QY 419 CCCCATATGCTACTGGAGATCAGTGGAGAAATTAATCCAGAGAAATGAAGAGATGGAGC 478
 Db 31305 CCCCATATGCTACTGGAGATCAGTGGAGAAATTAATCCAGAGAAATGAAGAGATGGAGC 31364
 QY 479 CAAGAGCAAAAGAAATACCCAGCTGTGGATGTGATGTGATATAAGCAAGTCCGATGCT 538
 Db 31365 CAAGAGCAAAAGAAATACCCAGC-----TGTGACTGTGTAGTAGAGCAAGGTCCAATGCT 31418

QY	539	GTAAAGAGCAATATTGTCATAGGAACCTGGAATGTCAGGTCCATGATCAAGGCAAAATTGG	598	QY	1589	-----	1588
DB	31419	GTAAAGAGCAATATTGTCATAGGAACCTGGAATGTCAGGTCCATGATCAAGGCAAAATTGG	31478	DB	32496	TCGCAATCAATGCAAGAAATAGAGGAAACAAACAGATGGGAAAAAATTAGGGATCTCT	32555
QY	599	AAGTGTCTCAAAACAGAGATGGCAAGAGTGAATGTCAACATTTCTAGGAATCAAGCAACTAA	658	QY	1589	-----	1588
DB	31479	AAGTGTCTCAAAACAGAGATGGCAAGAGTGAATGTCAACATTTCTAGGAATCAAGCAACTAA	31538	DB	32556	TCAAAAATCAGAGATACAAAGGAAACATTTCTATGCAAAAGATGGCTCGATAAAGACAGA	32615
QY	659	AATGCACTGGATGGTGAATTTAACTCAGATGACCAATTAATCTACTACTCGGGCAGG	718	QY	1589	-----	1588
DB	31539	AATGCACTGGATGGTGAATTTAACTCAGATGACCAATTAATCTACTACTCGGGCAGG	31598	DB	32616	AATGGTATGGACCTAAACAGAGCAGAGATATTAAAGAGAGATGGCAAGATACACAGAA	32675
QY	719	AATCCCTCAGAGAAATGGAGTAGCCATCATGTCAACAAAAGAGTCCGAAATCAGATAC	778	QY	1589	-----	1588
DB	31599	AATCCCTCAGAGAAATGGAGTAGCCATCATGTCAACAAAAGAGTCCGAAATCAGATAC	31658	DB	32676	GAACTGTACAAAAAAGATCTTTCAGACCCAGATAATCAGATGGTGTGATCACTGACCTA	32735
QY	779	TTGGATGCAAGTCTCAAAAAACGACAGATGATCTCTGTTTGTTCGAAGCAACCAATTC	838	QY	1589	-----	1588
DB	31659	TTGGATGCAAGTCTCAAAAAACGACAGATGATCTCTGTTTGTTCGAAGCAACCAATTC	31718	DB	32736	GATCCAGACATCTCTGGAAATGGAAGTCAAGTGGGCTTTAGAAAGCATCACTACGAACAA	32795
QY	839	ATATCAGATTAATCCAGTCTATGCCCCACCAAGTAACTGTAAGAGCTGAAGTTGAAC	898	QY	1589	-----	1588
DB	31719	ATATCAGATTAATCCAGTCTATGCCCCACCAAGTAACTGTAAGAGCTGAAGTTGAAC	31778	DB	32796	GCTAGTGGAGGTGATGGAATTCAGATTGAGCTATTTTAAATCCTGAAGTATGATGCTGTG	32855
QY	899	GGTCTATGAAGACCTACAGAGACCTTTTAGAACTAACACCCCAAAAAAGATGCTCTTCTCA	958	QY	1589	-----	1588
DB	31779	GGTCTATGAAGACCTACAGAGACCTTTTAGAACTAACACCCCAAAAAAGATGCTCTTCTCA	31838	DB	32856	AAAGTGTGCACCTCAATATGCCAGCAATTTGGAAAACTCAGCAGTGGCCCAAGACTGG	32915
QY	959	TTATAGGGACTGGAATGCAAAAGTAGAAGCAAGAAACACCTGGAGTAAACAGGCAANT	1018	QY	1589	-----	1588
DB	31839	TTATAGGGACTGGAATGCAAAAGTAGAAGCAAGAAACACCTGGAGTAAACAGGCAANT	31898	DB	32916	AAAAGTCCGTTTTTCATTTCCAAATCCCAAAGAAAGCAATGCCAAAGATGTCTCAACTAC	32975
QY	1019	TTGGCCTTGAATACGGAATGAAGCAGGCGCAAGACTTAATAGAGTTTTTGCAGAAAAATG	1078	QY	1589	-----	1588
DB	31899	TTGGCCTTGAATATGAATATGAAGTAGAGGCGCAAGACTTAATAGAGTTTTTGCAGAAAAATG	31958	DB	32976	CGCACAGTTGCACCTCATCTCACACGCTAGTAAGATTAATGCTCAAAATTTCTCOAGCCAGG	33035
QY	1079	CACGTGTCATAGCAACACCCCTTTTCAACAAACACAGAGAACTCTACATGACAT	1138	QY	1589	-----GAACTTCCAGTTGTTTCAAGCTGGTTTTAGAAAACTCAGA	1627
DB	31959	CACGTGTCATAGCAACACCCCTTTTCAACAAACACAGAGAACTCTACATGACAT	32018	DB	33036	CTTCAGCAATATGTGAAACCGTGAACTTCATATGTTCAAGCTGGTTTTAGAAAAAGCAGA	33095
QY	1139	CACAGATGTCACACCGGAATCAGATGATATATTTCTTCGACGCAAAAGATGAGAA	1198	QY	1628	GGAACCCAGAGACCAAAATTTGCCAAACATCTCTGTATCTATGGAAGAAAGCAAGAGTTCCAG	1687
DB	32019	CACAGATGTCACACCGGAATCAGATGATATATTTCTTCGACGCAAAAGATGAGAA	32078	DB	33096	GGAACCCAGAGATCAAAATTTGCCAAACATCTCGCTGGATCATGGAAGAAAGCAAGAGTTCCAG	33155
QY	1199	GCTCTATACAGTCAGCAAAAAACAGACCCAGAGCTTACTGTGCTCAGATCATGAACCTC	1258	QY	1688	AAAAACATCTATTTCTGCTTTTATGACTATGCAGAAAGCCTTTGACTGTGGGGGTCACAAT	1747
DB	32079	GCTCTATACAGTCAGCAAAAAACAGACCCAGAGCTTACTGTGCTCAGATCATGAACCTC	32138	DB	33156	AAAAGCATCTATTTCTGCTTTTATGACTATGCAGAAAGCCTTTGACTGTGGATCACAAT	33215
QY	1259	TTATGCAAAATTCAGACTTAAATTTGAAGAAAGTAGGAAAAACCACTAGATCACTCAGGT	1318	QY	1748	AAACTGTGAAAAATTTCTGAAAGGGATGGGAATACCAGACCACCTGACCTGACTCTTGAAA	1807
DB	32139	TTATGCAAAATTCAGACTTAAATTTGAAGAAAGTAGGAAAAACCACTAGATCACTCAGGT	32198	DB	33216	AAACTGTGAAAAATTTCTGAAAGAGATGGGAATACCAGACCACCTGATCTGCCTCTTGAGA	33275
QY	1319	AAGACCTTAAATCCCAATCCCTTATGATTAATACAGTGAAGTAGAATAATAGATTTAAGGGCC	1378	QY	1808	AATTTGTATGCAAGTCCAGGAACCAACAGTATGAACTGGACATGGAACCAACAGACTGTTTC	1867
DB	32199	ATGACCTCAGTCAAAATCCCTTATGATTAATACAGTGAAGTAGAATAATAGATTTAAGGGCC	32258	DB	33276	AATTTGTATGCAAGTCCAGGAACCAACAGTATGAACTGGACATGGAACCAACAGACTGTTTC	33335
QY	1379	TAGATCTGATAGACAGAGTACCTTAATGAACTATGGAACAGAGTTTCAATGATGACAGG	1438	QY	1868	CAAGTAGGAAAAAGAGTTCGTCAAGCTGTATTTGTCAAGCTGCTTATTTAACTTATAT	1927
DB	32259	TAGATCTGATAGACAGAGTACCTTAATGAACTATGGAACAGAGTTTCAATGATGACAGG	32318	DB	33336	CAATATGGAAGAAAGAGTTCGTCAAGCTGTATTTGTCAAGCTGCTTATTTAACTTATAT	33395
QY	1439	AGACAGGATCGACACCAATCCCAATGGAAGAAAGTAATGCAAAAAAGCAAAATGGCTGTCTG	1498	QY	1928	GCAGAG-ACATCATGAGAAACCGCTGGGTGGAAAGCAACAGCTGGAAATCAAGATTGCC	1986
DB	32319	AGACAGGATCGACACCAATCCCAATGGAAGAAAGTAATGCAAAAAAGCAAAATGGCTGTCTG	32378	DB	33396	GCAGAGTACATCATGAGAAACCGCTGGAAAGCAACAGCTTGAATCAAGATTGCC	33455
QY	1499	GGGAGGCTTACAAATAGCTGTGAAAAAGAGAGAGTGAAGCAAAAGGAAAAAGGAAA	1558	QY	1987	GGGAGAAATAGCAATAACCTCAGATATGCAGATATACCACCTTTATGGCAAAAGTGA	2046
DB	32379	GGGAGGCTTACAAATAGCTGTGAAAAAGAGAGAGTGAAGCAAAAGGAAAAAGGAAA	32435	DB	33456	AGGAGAAAATCAATAACCTCAGATATGCAGATATGCACCCCTTATGGCAAAAGTGA	33515
QY	1559	GATAAAGCATCTGAAATCCAGAGTTCCAA-----	1588	QY	2047	GAGGAACTAAAAAGCCTCTTTGATGAAGGTGAAGAGGAGAGTGAAAAAGTTGGCTTAAAG	2106
DB	32436	GATATAAACATCTGAATGCAGAGTTCCAAAGATAGCAAGAGATGAAGAAAGCCTTCT	32495	DB	33516	GAGGAACTAAAGAGCCTCTTTGATGAAGGTGAAGGTGAAGGTGAAGGTGAAGGTGAAGGT	33575
				QY	2107	CTCAACATTCAGAAACAGAGATCATGGCATCTGGTCCCATCACTTCATGGGAATAGAT	2166

```

Db 33576 CTCACATTCAGAAATGAGATATATGCGATCCGTCCTCATCTTCATGGAAATAGAT 33635
Qy 2167 GGGGAAACAGTGGAAACAGTGTGACAGCTTTATTTTGGGGGGCTCAAAATCACTCCAG 2226
Db 33636 GGAGAAACAGTGGAAACAGTGTGACAGCTTTATTTTCTGGGCTCAAAATCACTACAG 33693
Qy 2227 ATGGTGACTGCGAGCATGAAATTAAGACACTTACTCTCTGGGAGAAAGTTA--ACCA 2284
Db 33694 ATGTTGACTGAGCCATGAAATTAAGACCCCTTACTCTTGAAGGAAAGTTATGACCA 33753
Qy 2285 ACCTAGATAGCATATGAAAGACGAGACATTAACCTTGCACAAAGCCCATCTAGTCA 2344
Db 33754 ACCTAGATAGCATATGAAAGACGAGACATTAACCTTGCACAAAGGTTTCATCTAGTCA 33813
Qy 2345 AGGCTATGTTTTTCCAGTGTCATGATGATGAGTGTGAGTGTGAGAAAGCTG 2404
Db 33814 AGGCTATGTTTTTCCAGTGTCATGATGATGAGTGTGAGTGTGAGAAAGGCTA 33873
Qy 2405 AGCACTGAAGAAATGATGCTTTTGAACCTGTGGTCTGGAGAAAGCTCTTGAGAGTCCCTT 2464
Db 33874 AACCCGAGAAATGATGCTTTTGAACCTGTGGTCTGGAGAAAGCTCTTGAGAGTCCCTT 33933
Qy 2465 GGACTGCAAGGAGATCCAAACAGTCCATTTCTGAAGGAGATCAGCCCTGGGATTTCTTGG 2524
Db 33934 GGACTGCAAGGAGATCCAAACAGTCCATTTCTGAAGGAGATCAGCCCTGGGATTTCTTGG 33993
Qy 2525 AAGGAATGATCCTAAAGCTGAATCTCAGTACTTTTGGCCACCTGATCAGAAAGAGCTGACT 2584
Db 33994 AAGGAATGATCCTAAAGCTGAATCTCAGTACTTTTGGCCACCTGATCAGAAAGAGTGTACT 34053
Qy 2585 CACTGGAAAGACCTGATGCTGGAGGATTTGGGGCAGGAGAGAGAGGGGACGACAGA 2644
Db 34054 CACTGGAAAGACCTGATGCTGGAGGATTTGAGGACAGAGAGAGGGGATGCGAGA 34113
Qy 2645 GGATGAGATGCTGGAGGATCAGTACGTCGAGCTGAGCTGCTGGTGAATCTCTGG 2704
Db 34114 GGATGAGATGCTGGAGGATCAGTACGTCGAGCTGAGCTGCTGGTGAATCTCTGG 34173
Qy 2705 AGTTGATGAGGACGAGGAGGCTGCTCTGGGCGATTCTATGGGGTACACAAAGATTGGA 2764
Db 34174 AGTTGATGAGGACGAGGAGGCTGCTCTGGGCGATTCTATGGGGTACACAAAGATTGGA 34232
Qy 2765 CACACTGACCACTGAACTGAAGTGA 2794
Db 34233 CATGACTGAGTGAATGATCTGATCTGATCT 34262

```

```

RESULT 11
AC140963
LOCUS AC140963 213549 bp DNA linear HTG 28-MAR-2003
DEFINITION Bos taurus clone RP42-525L10, WORKING DRAFT SEQUENCE, 7 ordered
pieces.
ACCESSION AC140963
VERSION AC140963.2 GI:29336215
KEYWORDS HTG; HTGS PHASE2; HTGS_DRAFT.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 213549)
AUTHORS Akhtar, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Carligian, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J.,
Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E.,
Laric, P., Lee-Lin, S.-O., Legaspi, R., Maduro, Q.L., Maduro, V.B.,
Margulies, E.H., Masiello, C., Maskeri, B., McDowell, J.,
Paguigan, C., Pearson, R., Portnoy, M.E., Prasad, A.,
Reddix-Dugue, N., Schandler, K., Schueler, M.G., Sison, C.,
Stantripop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L.,
Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.
NISC Comparative Sequencing Initiative

```

```

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

```

Unpublished
2 (bases 1 to 213549)
Green, E.D.
Direct Submission
Submitted (06-MAR-2003) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 213549)
Green, E.D.
Direct Submission
Submitted (28-MAR-2003) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
On Mar 28, 2003 this sequence version replaced gi:28866995.
----- Genome Center
Center: NIH Intramural Sequencing Center
Web site: http://www.nisc.nih.gov
Contact: nisc_zoo@hgr.nih.gov
----- Project Information
Center project name: djt
Center clone name: 525L10

```

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

```

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 212561 bases at least Q40
Consensus quality: 212813 bases at least Q30
Consensus quality: 212927 bases at least Q20
Insert size: 155000; agarose-fp
Insert size: 212949; sum-of-contigs
Quality coverage: 12.59x in Q20 bases; agarose-fp
Quality coverage: 9.17x in Q20 bases; sum-of-contigs
-----

```

* NOTE: This is a 'working draft' sequence. It currently consists of 7 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.

```

* 1 1801: contig of 1801 bp in length
* 1802 1901: gap of unknown length
* 41678 41777: contig of 39776 bp in length
* 41778 81698: contig of 39921 bp in length
* 81699 81798: gap of unknown length
* 81799 95185: contig of 13387 bp in length
* 95186 95285: gap of unknown length
* 95286 106702: contig of 11417 bp in length
* 106703 106802: gap of unknown length
* 106803 152491: contig of 45689 bp in length
* 152492 152591: gap of unknown length
* 152592 213549: contig of 60958 bp in length.

```

```

FEATURES
Location/Qualifiers
1..213549
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="RP42-525L10"

```

TITLE

```
misc_feature      /clone lib="RP42"
1..1801
/note="assembly_fragment
clone_end:T7
vector_side:left"
misc_feature      1902..41677
/note="assembly_fragment"
41778..81698
/note="assembly_fragment"
81799..95185
/note="assembly_fragment"
95286..106702
/note="assembly_fragment"
106803..152491
/note="assembly_fragment"
129729..213549
/note="clone overlaps with GenBank Accession Number
AC140962 clone RP42-329016 (center project name dju)"
152592..213549
/note="assembly_fragment
clone_end:SP6
vector_side:right"

Query Match      59.0%; Score 2107.7; DB 1; Length 213549;
Best Local Similarity 86.1%; Pred. No. 0;
Matches 2504; Conservative 0; Mismatches 243; Indels 161; Gaps 17;

Qy 22 GAGAGTACCTACCTCGTCCAGGTAAGGAGCAGTAGCTGCGTTTGCTGGAGAGCGCG 81
Db 24602 GAGAGGAGATACCCCGCTCCAGGTAAGGAGCAGCGCGCTTTGCTGGAGAGCGCC 24661

Qy 82 -TAAAGAGATACCCCGCCAGGTAAGGAGAAACCCCAAGTAAGATGCTAGGTTGTGA 140
Db 24662 ATGAAGAGACACCACTGTCAGATAGAGAAACCCCACTAAGACGTTAGGTTGTGA 24721

Qy 141 GAGGCGATCAGAGGCGCAGACATCTGAAACCATACACGCGAGAAACCTAGTCAATCTAATC 200
Db 24722 CAGGCGATCAGAGGCGCAG---ACTGAAACCACAATCACAGAAACCTAGCCAATCTAATC 24777

Qy 201 ACATAGAGACACGAGCTTGTCTAATCAATGAATCAAGCCATGCC-CGTGGGCAACC 259
Db 24778 ACA--TGGACCATAGCTTATCTAATCAATGAATCAAGCCATGCCATATAGAGCCACC 24835

Qy 260 CAAGATGGCAGGTCATGGTGGAGAGATCTGACAGAAATGGTCCACTGGAGAGGGAAT 319
Db 24836 CAAGACGTCAGGTCACGGTGGAGAAATCTGACAAATTTGGTCCACTGGAGAGGGAAT 24895

Qy 320 -GCAAAACACATTCATGATCTTTGCTTGGAGAACCCCATGAACAGTATGAAAGGCAAAAT 378
Db 24896 GGCAAAACCACTTCAATATTTCTTGCTTGGAAATCCCATGAATAGTATGAAAGGCAAAA 24955

Qy 379 GATAGGATCTGAAGAGGAACTCCCGAGGTCAGTGGTCCCATATGCTACTGGAGAT 438
Db 24956 GATAGGACACTGAAGATGAATCTCCCGAGTGGACAGTGGCCCATATGCTACTGGAGAT 25015

Qy 439 CAGTGGAGAAATAACTCCAGAAAGAAATGAAGAGATGGAGCCAAAGCAAAA----- 488
Db 25016 CAGTGGAGAAATAACTCCAGACAGAAATGAAGAGCAGAGCCAAAGCAAAAACAAA 25075

Qy 489 ---AGAAATACCCAGCTGGATGTGACTGGTGATATGAAGCAGGTCGATGCTGAAGA 545
Db 25076 AACAAAACCCAGTGTGCAATGTGACTGGTGAGAGAGCAGGTCGAATGCTGTAAAGA 25135

Qy 546 GCAATATTGCATAGGAACCTGGAATGTGAGTCCATGAATCAAGGCAAAATTTGGAAGTGGT 605
Db 25136 GCAATATTGCATAGGAACCTGGAATGTGAGTCCATGAATCAAGGAAATTTGGAAGTGG- 25194

Qy 606 CAACAGAGATGGCAAGAGTGAATGTCAACATTTCTAGGAATCAGCGCACTAAATGGAC 665
Db 25195 CAATAGAGATGGCAAGAGTGAACATCAACATTTCTAGGAATCAG-TAACTAAGATGGAC 25253

Qy 666 TGGATGGGTGAATTTAACTCAGATGACCATATATATCTACTCTACCTGCGGCGAGGAATCCCT 725
```

```
Db 25254 TGCATGGGTGAACCTTAATCAGATGACCATATATCTACTATATAGTGGGCAGAATCCCT 25313
Qy 726 CAGAAGAAATGGAGTAGCCATCATGGTCAACAAAAGAGTCGAAATGCGACTTGGATG 785
Db 25314 TAGAAGAAATTAATAGCCATCATAGTCAACACAGAAATCAAAATGCACTACTTGGATG 25373
Qy 786 CAGTCTCAAAAACGACAGATGATCTCTGTTGTTTCCAGGCAAAACCAATCAATATCAC 845
Db 25374 TAATCCC-AAATGACAGATGATCCTGGTCTGTTTCCAGGCAAAACCAATCAATATCAC 25432
Qy 846 AGTAATCCAAGTCTATGCCCAACCCAGTAATGCTGAAGAGCTGAAGTTGAACGGTCTCA 905
Db 25433 AGTAATCCAAGCTATGCCCAACCCAGTAACGCTGAAGAGCTGAAGTTGAATGGTCTG 25492
Qy 906 TGAAGACCTACAGACCTTTTAGAAGTCAACCCCAAAAGAGTCTCTCTCATATAGG 965
Db 25493 TGAAGACCTACAGACCTCTTTAGAACCTAACCCCTAAAAAAGATGCTCTTTCAITATAGG 25552
Qy 966 GGACTGGAAATGCMAAAGTAGGAAGCAAAAGAACACCTCGAGTAAACAGGCAAAATTTGGCT 1025
Db 25553 GAACTGGAAATGCMAAAGTGGGAACCTCAAGAAACACCTCGAGTAAACAGGCAAAATTTGGCT 25612
Qy 1026 TGAATAGCGAATGAAGCAGGCAAAAGACTAATAGAGTTTTGCCAAGAAATGCACTGGT 1085
Db 25613 TGAATAGCAGAAATGAAGCAAGGCAAAAGGCTAATAGAGTTTTGCCAAGAAATGCACTGGT 25672
Qy 1086 CATAGCAACACCCCTTCCCAACACACAGAGAGAGCTCTACACATGCAATCACCAAGA 1145
Db 25673 CATAGCAACACCCCTTCTCCCAACACAC-----AAGGCTCTACATGAAACATCACCAAGA 25727
Qy 1146 TGGTCAACACCCAAATCAGATTGATTTATTTTTCAGGCAAAAGATGGAGAGCTCTAT 1205
Db 25728 TGCCCAACACCCAAATCAGATTGACTATATTTCTTTCAGGCAAA--TTGAGAAAGCTCTAC 25785
Qy 1206 ACAGTCAGCAAAAACAGACAGGAGGCTTACTGTGGCTCAGATCATGACTGCTTATGCG 1265
Db 25786 AAAGTCAGCAAAAACAGACAGGAGGCTGCTGACTCAGATCATGAACT--CTTATGCG 25844
Qy 1266 CAAATTCAGACTTAAATGAAGAAATAGGGAACCCACTAGATCACTCAGTAAGAGCT 1325
Db 25845 CAAATTCAGACTTAAATGAGGAAGTAGGGAACCCACTAGACCATCAAGAAATGACCT 25904
Qy 1326 AAATCAATTCCTTATGATTTATACAGTGGAAATGAGAAATAGATTTAAGGCTTATGCT 1385
Db 25905 AAATCAATTCCTTATGATTTATACAGTGGAAATGCAATAGATTTAAGGCTTATGCT 25964
Qy 1386 GATAGCAGAGTACCTAATGAAGTATGACAGAGGTTTATGACATTTGACAGGACAGG 1445
Db 25965 GATAGCAGAGTGCCTTAATGAAGTATGACAGAGTTTATGACATTTGACAGGACAGG 26024
Qy 1446 GATCGAGACCATCCCATGGAAGAAAGATGCAAAAGAAAGCAAAATGGCTGCTTGGGGAGCG 1505
Db 26025 GATCAAGACCATCCCATGGAAGAAAGAT-----GCAAAATGGCTGCTTGAAGAGCG 26076
Qy 1506 CTTACAAATAGCTGTGAAGAGAGAGAGTGAAGAGCAAGGAAAGGAAAGGAAAGATATAA 1565
Db 26077 CTTACAAATAGCTGTGAAGAGAGAGAGTCAAAAGCAAGGAAAGGAAAGATATAAT 26136
Qy 1566 GCATCTGAATGCAAGTCTTCCAAA----- 1588
Db 26137 GCATTTGTTTTCATCCAAACCCCAAGAAAGCAATGCCAAAGTATATCTCAAACTACCGC 26196
Qy 1589 ----- 1588
Db 26197 ACAATTCGACTCATCTCACACGCTAGTAAGTAAATGCTCAAAATTTCTCAAGCCAGGCTT 26256
Qy 1589 -----GAACTTCAGTTGTTCAAGCTGGTTTGAAGAAAGTCAAGAGA 1630
Db 26257 CAGCAATATGCGAACCGTGAACCTTCTGATGTTTCAAGCTGGTTTGAAGAAAGCAGAGA 26316
Qy 1631 ACCAGAGACCAATTCGCAACATCTCTGTATCATGTGAAAAAAGCAAGAGAGTTCCAGAAA 1690
Db 26317 ACCAGATATCAATTCGCAATATCCACTGGAACCATGGAAGGAAAGCAAGAGTTCCAGAAA 26376
```

QY 1691 AACATCTATTTCTGCTTTTATTGACTATCAAAAGCCTTTGACTGTGGGGGTCAATAAAA 1750
Db 26377 AACATCTATTTCTGCTTTTATTGACTATGCCAAGCCTTTGACTGTGGGTCAATAAAA 26436
QY 1751 CTGTGGAATTTCTGAAGGATGGGAATACAGACACACCTGACCTGACTCTTGAATAAT 1810
Db 26437 CTGTGGAATTTCTGAAGGATGGGAATACAGACACACCTGACTCTGCTCTTGAAGAAC 26496
QY 1811 TTGTATGCAAGTTCAGGAAGCAACAGTTAGAACTGCAATGGAACAAACAGACTGCTTCCAA 1870
Db 26497 TTGTATGTAGTTCAGGAAGCAACAGTTAGAACTGCAATGGAACAAACAGACTGCTTCCAA 26556
QY 1871 GTAGAAAAGAGATGTCAGAGCTGTATATGTGCAACCCGCTTGTAACTTCTATGCA 1930
Db 26557 ATAGAAAAGAGATTCGTCAAGGCTGTATATGTGCAACCCGCTTGTAACTTCTATGCA 26616
QY 1931 GAG-ACATCATGAGAAACGCTGGCTGGAAGACACAGCTGCAATCAAGATTGCCGG 1989
Db 26617 GAGTACATCATGAGAAACGCTGGCTGGAAGACACAGCTGCAATCAAGATTGCCGG 26676
QY 1990 AGAATAGCAATTAACCTCAGATATGCAATGATACACCCCTTATGCGAGAAAGTGAAG 2049
Db 26677 AGAATATCAATTAACCTCAGATATGCAATGATACACCCCTTATGCGAGAAAGTGAAG 26736
QY 2050 GAACATAAAGCCTCTTATGCAAGTGTGAAGAGAGAGTGAAGAGTTCCTTAAAGCTC 2109
Db 26737 GAACATAAAGCCTCTTATGCAAGTGTGAAGAGAGAGTGAAGAGTTCCTTAAAGCTC 26796
QY 2110 AACATTGAGAAACGAGATCATGCGATCTGCTGCCATCACTTCATGGGAAATAGATGG 2169
Db 26797 AACATTGAGAAACGAGATCATGCGATCTGCTGCCATCACTTCATGGGAAATAGATGG 26856
QY 2170 GAACAGTGGAAACAGTGTGAGCTTATTTTGGGGGGCTCCAAATCACTGCGAGTG 2229
Db 26857 GAACAGTGGAAACAGTGTGAGCTTATTTTGGGGGGCTCCAAATCACTGCGAGTG 26914
QY 2230 GTGACTGAGCAGCATGAATTAAGACACATTTACTCTTGGGAAGAAAGTTA--ACCAAC 2287
Db 26915 GTGACTGAGCAGCATGAATTAAGACACATTTACTCTTGGGAAGAAAGTTAAGCAAC 26974
QY 2288 TAGATAGCATATTGAAAGCAGAGACATTTACTCTTGCACAAAGCCCATCTAGTCAAG 2347
Db 26975 TAGATAGCATATTGAAAGCAGAGACATTTACTCTTGCACAAAGCTTCGTAGTCAAG 27034
QY 2348 CTATGTTTTTCCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2407
Db 27035 CTATGTTTTTCCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 27094
QY 2408 ACTGAAGATTTGATGCTTTTGAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2467
Db 27095 GCCAAAGATTTGATTTTGAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 27154
QY 2468 CTGCAAGAGATCCAAACAGTCTTCTGAAGGATGACCCCTGGGATTTCTTGAAG 2527
Db 27155 CTGCAAGAGATCCAAACAGTCTTCTGAAGGATGACCCCTGGGATTTCTTGAAG 27214
QY 2528 GAATGATGCTAAAGCTGAACTCCAGTACTTTGGCCACCTGATCAGAAAGCTGACTCAC 2587
Db 27215 GAATGATGCTAAAGCTGAACTCCAGTACTTTGGCCACCTGATGAGAGTTGATCAT 27274
QY 2588 TGGAAAAGACCTGATGCTGGAGGATTTGGGGCAGGAGAGAGGGGACGACAGGA 2647
Db 27275 TGGAAAAGACCTGATGCTGGAGGATTTGGGGCAGGAGAGAGGGGACGACAGGA 27334
QY 2648 TGAGATGCTGATGGGATCACTCACTCGATGAGCTGAGTCTGGGTGAACCTCTGAGT 2707
Db 27335 TGAGATGCTGATGGGATCACTCACTCGATGAGTCTGAGTCTGAGTGAACCTCAAGTGT 27394
QY 2708 TGGTATGAGACAGGAGGCTCTCTCGCGCATTTATGCGGTGACAAAGAGTTGACAC 2767
Db 27395 TGGTATGAGACAGGAGGCTCTCTCGCATTTATGCGGTGACAAAGAGTTGACAC 27454

QY 2768 GACTGAGCAACTCAACTGAACCTGAACCTG 2795
Db 27455 GACTGAGCAACTGATCTGATTGATCTG 27482

RESULT 12
AC144887/c
LOCUS

DEFINITION
AC144887 196994 bp DNA linear HTG 18-JUN-2003
Bos taurus clone RP42-433E9, WORKING DRAFT SEQUENCE, 8 ordered
pieces.

ACCESSION
AC144887 GI:31880107
VERSION
HTG; HTGS_PHASE2; HTGS_DRAFT.

KEYWORDS
SOURCE
ORGANISM

Bos taurus
Bos taurus (cow)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 196994)

REFERENCE
AUTHORS

Antonellis A., Ayele K., Beckstrom-Sternberg S.M., Benjamin B.,
Blakesley R.W., Bouffard G.G., Brinkley C., Brooks S., Cariaga K.,
Chu G., Coleman B., Coleman H., Engle J., Granite S., Guan X.,
Gupta J., Haghighi P., Han J., Hansen N., Ho S.-L., Hu P.,
Hurle B., Idol J.R., Karlins E., Kwong P., Laric P., Lee-Lin S.-Q.,
Legaspi R., Maduro Q.L., Maduro V.B., Margulies E.H., Masiello C.,
Maskeri B., McQuell J., Peguirigan C., Pearson R., Portnoy M.E.,
Prasad A., Reddi-Dugue N., Schandler K., Schueler M.G., Shah K.,
Simon C., Stantrapop S., Thomas J.W., Thomas P.J., Taipouri V.,
Vogt J.L., Wetherby K.D., Wiggins L., Young A. and Green E.D.
NISC Comparative Sequencing Initiative

TITLE
JOURNAL

REFERENCE
2 (bases 1 to 196994)

Direct Submission

Submitted (23-MAY-2003) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA

REFERENCE
3 (bases 1 to 196994)

Green E.D.

Direct Submission

Submitted (18-JUN-2003) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA

On Jun 18, 2003 this sequence version replaced gi:31044307.

----- Genome Center

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: <http://www.nisc.nih.gov>

Contact: nisc_zooenhgr@nisc.nih.gov

----- Project Information

Center project name: eeg

Center clone name: 433E09

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 195487 bases at least Q40
Consensus quality: 195921 bases at least Q30
Consensus quality: 196170 bases at least Q20
Insert size: 167000; agarose-fp
Insert size: 196294; sum-of-ctgigs
Quality coverage: 12.61x in Q20 bases; agarose-fp

QY 1338 TTATGATTATACAGTGGAGTGCAGAAATAGATTTAAGGCGCTAGATCTGTATAGACAGAGT 1397
Db 50791 TTATGATTATACAGTGGAGTGCAGAAATAGATTTAAGGCGCTAGATCTGTATAGATAGAGT 50732
QY 1398 ACCTAATGAACATATGGACAGAGGTTTCATGACATTTGTACAGAGACAGAGGATCGAGACCAT 1457
Db 50731 GCCTGATGAACATATGGAAATGAGGTTTCGTGACATTTGTACAGAGACAGAGATCAAGACCAT 50672
QY 1458 CCCATGAAAGAAATCCAAAGAAAGCAAAATGCTGTCTGGGAGCGCTTACAAATAGC 1517
Db 50671 CCCATGAAAGAAATCCAAAGAAAGCAAAATGCTTCTGGGAGCGCTTACAAATAGC 50612
QY 1518 TGTGAAAGAAAGAGAGTGAAGAAAGCAAGGAAAGAAAGAAAGATATAAAGCATCTGAAATGC 1577
Db 50611 TGTGAAAGAAAGAGAGTGAAGAAAGCAAGGAAAGAAAGAAAGATATAAAGCATCTGAAATGC 50552
QY 1578 AGAGTTCCAAA----- 1588
Db 50551 AGAGTTCCAAAAGATAGCAAGAGAGATAGAAAGACATTTCTTCAGCGATCAGTGCRAAGA 50492
QY 1589 ----- 1588
Db 50491 ATTAGAGAAACCAAGATGGAAAGACTAGAGATCTCTTCAAGAAATCAGAGATAC 50432
QY 1589 ----- 1588
Db 50431 CAAAGGAACATTTATGCAAGGATGAGTCTGATTAAGGACAGAAATGGTATGGACCTAAC 50372
QY 1589 ----- 1588
Db 50371 AGAGCAGAGATATTAAAGAGATGGCAAGATACACAGAGAACTGTACAAAAG 50312
QY 1589 ----- 1588
Db 50311 ATCTTCAGCCAGATAATCAGATGGTGTGATCACTACCTAGAGCCAGACATCTCG 50252
QY 1589 ----- 1588
Db 50251 AATGTGAAGTCAAGTGGCGCTTCAAGAGCATCTATGAACTAAGTAGTGGAGGTGATA 50192
QY 1589 ----- 1588
Db 50191 GAATTCAGTTGAGCTATTCCAAATTTCTGAAGATGATGCTGTTAAAGTCTGCATCAA 50132
QY 1589 ----- 1588
Db 50131 TATGCCAGCAAAATTGGAAAACTCAGCAGTGGCCACAGGACTGGACAAAGGTCAAGTTTCA 50072
QY 1589 ----- 1588
Db 50071 TTCCAATCTCAAGAAAGCAATGCCAAGAAATGCTCAAACTACCCCAAAATGCACTCA 50012
QY 1589 -----GAACTTT 1594
Db 50011 TCTCACATGCTAGTAAAGTAATGCTCAAAATTTCTCAAGCCAGGCTTCAGCAATGAATTT 49952
QY 1595 CCAGTTGTTCAAGCTGGTTTACAAAGTGCAGAGGAAACAGAGACCAAAATTCGCCAATC 1654
Db 49951 CCAGTTGTTCAAGCTGGTTTACAAAGGCGAGAGAACAGAGATCAAAATTCGCCAATC 49892
QY 1655 CTCGTATCATGAAAGAAAGCAAGAGAGTTCCAGAAAGAAATCTATTTCTGCTTTATTGAC 1714
Db 49891 CGCTGATCATGAAAGAAAGCAAGAGATTTCCAGAAAGAAATCTATTTCTGCTTTATTGAC 49832
QY 1715 TATGCAAAAGCCCTTTCAGCTGTGGGCTCACAATTAACCTGTGAAATTTCTGAAAGGATG 1774
Db 49831 TATGCCAAAGCCCTTTCAGCTGTGGGATCACAATTAACCTGTGAAATTTCTGAAAGGATG 49772
QY 1775 GGAATACCAAGACCTGACCTGACTCTTGAAAGAAATTTGTATGCAAGGTGAGGAGCAACA 1834
Db 49771 GGAATACCAAGACCTGACCTGACTCTTGAGAAATCTACAT--AGGCCAGGAGCAACA 49714
QY 1835 GTTAGAACTGGACATGGAACACAGACTGGTTCCAAAGTAGGAAAGGAGATATGTCAAGGC 1894

Db 49713 GTTAGAACTGGACATGGAACAAAGAGCTGGTTCCAAATAGGAAAGGAGCAGCTCAAGGC 49654
QY 1895 TGTATATTCTCACCCCGCTTGTAACTTCTATGACAGAG-ACATCATGAGAAACCGCTGGG 1953
Db 49653 TGTATATTATCACTCTGTTTATTAACTTATATGACAGATACATCATGAGAAACCGCTGGA 49594
QY 1954 CTGGAAGAGACACAGCTGGAATCAAGATTGCGGGAGAAATAGCAATAACCTCAGATAT 2013
Db 49593 CTGGAAGAAACCAAGAGCTGGAATCAAGATTGCTGGAGAAATATCAATAACCTCAGATAT 49534
QY 2014 GCAGATGATACCAACCTTATGCGAGAAAGTGAAGAGGAACTAAAAAGCCTCTTGATGAAG 2073
Db 49533 GCAGATGACACCACTTATGCGAGAAAGTGAAGAGAACTAAAGAGCCTCTTGATGAAG 49474
QY 2074 GTGAAAGAGAGAGTGAAGAAAGTGTGGCTTAAAGCTCAACATTCAGAAACCGAAGATCATG 2133
Db 49473 GTGAAAGAGAGAGTGAAGAAAGTGTGGCTTAGAGCTCAACATTCAGAAATGAAGATCATG 49414
QY 2134 GCATCTGTGCTCCATCCTCATGCGAATAGTGGGAAACAGTGGAAACAGTGTGACAG 2193
Db 49413 GCATCTGTGCTCCATCCTCATGCGAATAGTGGGAAACAGTGGAAACAGTGTGACAG 49354
QY 2194 TTTATTTTGTGGGGGCTCCAAATCACTGCAGATGCTGAGCTGCAGCCATGAAATTTAAA 2253
Db 49353 TTTATTTT--CTGGGCTCCAAATCACTGCAGATGCTGATTCAGCCATGAAATTTAAA 49296
QY 2254 GACACTACTCTCTGGAGAAAGTTA--ACCACTTAGATAGATATTGAAAGACAGAG 2311
Db 49295 GATGCTTACTCTTGGAAAGAAAGTTATGACCAACCTAGATAGCATATTCAAAAGACAGAG 49236
QY 2312 ACATTAACCTTCGCAACAAAGCCCATCTAGTCAAGGCTATGTTTTCCTGCTGCTATGT 2371
Db 49235 ACATTAACCTTCGCAACAAAGGTTCTGCTAGTCAAGGCTATGTTTTCCTGCTGCTATGT 49176
QY 2372 ATGATGTGAGAGTTGACTGTGAAGAAAGCTGAGCACTGAAGAAATGATGCTTTTGAAC 2431
Db 49175 ATGATGTGAGAGTTGGA-TGTGAAGAAAGCTGAGCGCCGAAAGAAATGATGCTTTTGAAC 49117
QY 2432 TGTGTGTGTGAGAGAGACTCTTGAAGTCCCTTGCAGTCCCAAGAGATCCAAACAGTCCA 2491
Db 49116 TGTGTGTGTG-----GAGTCCCTTGCAGAGAGAGATCCACCAGTCCA 49071
QY 2492 TTCTGAAGAGATCAGCCCTGGGATTTCTTTGGAAGAAATGATGCTAAAGCTGAAACTCC 2551
Db 49070 TTCTGAAGAGATCAGCCCTGGGATTTCTTTGGAAGAAATGATGCTAAAGCTGAAACTCC 49011
QY 2552 AGTACTTTGGCCACTGATCAGAGAGCTGACTCACTGGAAGAACCCCTGATGCTGGAG 2611
Db 49010 AGTACTTTGGCCACTCATGCAAGAGCTGACTCACTTTGGAAGAAAGACTCTGATGCTGGAG 48951
QY 2612 GGATTTGGGGCAGAGAGAGAGGAGGACAGAGAGATGAGATGGCTGGATGGCATCACTG 2671
Db 48950 GGATTTGGGGCAGAGAGAGAGGAGGAGGACAGAGAGATGAGATGGCTGGATGGCATCACTG 48891
QY 2672 ACTCGATGAGAGTGTGCTGGGTGAACCTCTGAGTGTGATGAGACAGGAGGCTGTC 2731
Db 48890 ACT-GATGGAGCTGAGTCTCAGTGAACCTCCGGGAGTTGGTGTGATGGACAGGAGGCTGTC 48832
QY 2732 CTGGGGGATTCATGGGCTCAAAAGATTGGACACAGCTGAGCACTGAACTGAACCTGA 2791
Db 48831 ATGCTGCGATTCATGGGCTCACAAGAGTGGACAGCTGAGCGGACTGAACTGAACCTGA 48772
QY 2792 ACTGT 2796
Db 48771 AAGGT 48767

RESULT 13
AC139312/cLOCUS AC139312 192199 bp DNA linear HTG 21-FEB-2003
DEFINITION Bos taurus clone RP42-482K4, WORKING DRAFT SEQUENCE, 10 ordered
pieces.


```

ACCESSION      AC139312
VERSION        AC139312.2  GI:28460773
KEYWORDS       HTG; HTGS PHASE2; HTGS_DRAFT.
SOURCE         Bos taurus (cow)
ORGANISM       Bos taurus

REFERENCE
AUTHORS        Akhter,N., Antonellis,A., Ayele,K., Beckstron-Sternberg,S.M.,
               Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
               Cariaga,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J.,
               Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,B.,
               Lalic,P., Lee-Jin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,
               Marquiles,E.H., Masello,C., Maskeri,B., McDowell,J.,
               Paquirigan,C., Pearson,R., Portnoy,M.E., Prasad,A.,
               Reddix-Dugue,N., Schandler,K., Schueler,M.G., Sison,C.,
               Skantripo,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.L.,
               Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
TITLE          NISC Comparative Sequencing Initiative
JOURNAL        Unpublished
REFERENCE      2 (bases 1 to 192199)
AUTHORS        Green,E.D.
TITLE          Direct Submission
JOURNAL        Submitted (30-JAN-2003) NIH Intramural Sequencing Center, 8717
               Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE      3 (bases 1 to 192199)
AUTHORS        Green,E.D.
TITLE          Direct Submission
JOURNAL        Submitted (21-REB-2003) NIH Intramural Sequencing Center, 8717
               Grovemont Circle, Gaithersburg, MD 20877, USA
COMMENT        On Feb 21, 2003 this sequence version replaced gi:28144407.
               ----- Genome Center
               Center: NIH Intramural Sequencing Center
               Center code: NISC
               Web site: http://www.nisc.nih.gov
               Contact: nisc.zoo@hgrl.nih.gov
               ----- Project Information
               Center project name: ddk
               Center clone name: 482K04

The sequence data in this record represents an 'enhanced'
version of a Phase 2 submission. Specifically, the indicated
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8X average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associated
with a Phrap-derived quality score.
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 190322 bases at least Q40
Consensus quality: 190839 bases at least Q30
Consensus quality: 191112 bases at least Q20
Insert size: 160000; agarose-fp
Insert size: 191299; sum-of-contigs
Quality coverage: 10.93x in Q20 bases; agarose-fp
Quality coverage: 9.14x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and

```

```

* the accession number will be preserved.
* 1 74471: contig of 74471 bp in length
* 74472 74571: gap of unknown length
* 74572 79290: contig of 4719 bp in length
* 79291 79390: gap of unknown length
* 79391 85712: contig of 6322 bp in length
* 85713 85812: gap of unknown length
* 85813 92766: contig of 6954 bp in length
* 92767 92866: gap of unknown length
* 92867 109583: contig of 16717 bp in length
* 109584 109684: gap of unknown length
* 109684 134062: contig of 24379 bp in length
* 134063 134163: gap of unknown length
* 134163 144366: contig of 10204 bp in length
* 144367 144466: gap of unknown length
* 144467 148230: contig of 3764 bp in length
* 148231 148330: gap of unknown length
* 148331 159617: contig of 11287 bp in length
* 159618 159717: gap of unknown length
* 159718 192199: contig of 32482 bp in length.
FEATURES
source
1..192199
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="RP42-482K4"
/clone_lib="RP42"
1..74471
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
misc_feature
1..86488
/note="clone overlaps with GenBank Accession Number
AC139732 clone RP42-145M1 (center project name ddj)"
misc_feature
74572..79290
/note="assembly_fragment"
79391..85712
/note="assembly_fragment"
/note="assembly_fragment"
85813..92766
/note="assembly_fragment"
92867..109583
/note="assembly_fragment"
106684..192199
/note="clone overlaps with GenBank Accession Number
AC139632 clone RP42-18C2 (center project name ddl)"
misc_feature
109684..134062
/note="assembly_fragment"
misc_feature
134163..144366
/note="assembly_fragment"
misc_feature
144467..148230
/note="assembly_fragment"
misc_feature
148331..159617
/note="assembly_fragment"
159718..192199
/note="assembly_fragment"
clone_end:T7
vector_side:right"
Query Match 58.9%; Score 2105.2; DB 1; Length 192199;
Best Local Similarity 77.2%; Pred. No. 0;
Matches 2581; Conservative 0; Mismatches 153; Indels 609; Gaps 8;
QY 51 GAGCAGTAGTCGGCTTTGCTGGAGCAGCGGTAAAGAGATACCCACGCCCAAGGTAAGA 110
DB 151713 GAGCATGGCTGTGCTTTGCTGGAGCAGCCATGAAGATACCCACGCCCAAGGTAAGA 151654
QY 111 GAAACCCAGTAAAGATGGTAGTGTGTGAGAGGCGCATCAGAGGCGACACATCTGAAC 170
DB 151653 GAAATCCAAAGTAAAGACGGTAGCTGTGCAAGAGGCGCATCAGAGGCGCAACACACCTGAACC 151594
QY 171 CATACCGCAGAAAACACTAGTCAATCAATCACTAGGACACACGCTTGCTACTCAA 230
DB 151593 CATATCCAGAAAACACTAGTCAATCAATCACTAGGACACACGCTTGCTGAACCTCAA 151534

```

QY 231 TGAAACTAAGCCATGCCCGTGGGCAACCCCAAGATGGGCAAGGTCAATGGTGGAGAGATCTG 290
Db 151533 TGAAACTAAGCCATGCCCGTGGGCAACCCCAAGATGGGCAAGGTCAATGGTGGAGAGATCTG 151474
QY 291 ACAGAACTGTGTCCACTGGAGAAAGGGAAT-GCAAACCACTTTCAGTATCTCTTGGCTTGAGA 349
Db 151473 ACAGAACTGTGTCCACTGGAGAAAGGGAATGGCAAACCACTTTCAGTATCTCTTGGCTTGAGA 151414
QY 350 ACCCATGAACAGTATGAAAGGCAAAATGATAGGATCTGAAAGAGAACTCCCCAGGT 409
Db 151413 ACCCATGAACAGTATGAAAGGCAAAATGATAGGATCTGAAAGAGAACTCCCCAGGT 151354
QY 410 CAGTAGTGCCCAATATGCTACTCGAGATCAGTGGAGAAATTAATCTCCAGAAAGAACTGAAG 469
Db 151353 CAGTAGTGCCCAATATGCTACTCGAGATCAGTGGAGAAATTAATCTCCAGAAAGAACTGAAG 151294
QY 470 AGATGGAGCCCAAGCAAAAGAAATACCCAGCTGTGGATGTGACTGGTGATATAGCAAGG 529
Db 151293 GGATGGAGCCCAAGCAAAAGAAATTAATCTCAGCTGTGGATGTGACTGGTGATAGCAAGG 151234
QY 530 TCCGATGCTGTAAGAGCAATATTGCAATAGGAACCTGGAACTGTCAGTCCATGNAATCAAG 589
Db 151233 TCCGATGCTGTAAGAGCAATATTGCAATAGGAACCTGGAACTGTCAGTCCATGNAATCAAG 151174
QY 590 GCAAAATTGGAAGTGGTCAACCAAGAGATGGCAAGAGTGAATGTCAACATTTCTAGGAATCA 649
Db 151173 GCAAAATTGGAAGTGGTCAACCAAGAGATGGCAAGAGTGAATGTCAACATTTCTAGGAATCA 151114
QY 650 GCGAACTAAATGGACTGGAATGGTGAATTTAACTCAGATGACCAATATATCTACTACT 709
Db 151113 GTGAATGAAATGGACTGGAATGGTGAATTTAACTCAGATGACCAATATATCTACTACT 151054
QY 710 GCGGCGAGGAATCCCTCAGAGAAATGGAGTAGCCATCATGGTCAACAAAGAGTCCGAA 769
Db 151053 GCGGCGAGGAATCCCTCAGAGAAATGGAGTAGCCATCATGGTCAACAAAGAGTCTGAA 150994
QY 770 ATGCACTACTTGGATGCACTCTCAAAACGACAGATGATCTCTGTTTGTTCCTCAAGGCA 839
Db 150993 ATGCACTACTTGGATGCACTCTCAAAACGACAGATGATCTCTGTTTGTTCCTCAAGGCA 150934
QY 830 AACCATCAATATCACAGTATCAAGTCTATGCCCAACCAAGTAACTGCTGAGAGAGCTG 889
Db 150933 AACCATCAATATCACAGTATCAAGTCTATGCCCAACCAAGTAACTGCTGAGAGAGCTG 150874
QY 890 AAGTTGAACCGTCTCTATGAAGACCTAAGAGACCTTTTGAAGTAAACACCCCAAAAGATG 949
Db 150873 AAGTTGAACCGGCTCTATGAAGACCTAAGAGACCTTTTGAAGTAAACACCTCAAAAGATG 150814
QY 950 TCCTTCTCATTTATAGGGAATGGAATGCAAAAGTAGGAAGCAAGAAACACCTGGAGTAA 1009
Db 150813 TCCTTCTCATTTATAGGGAATGGAATGCAAAAGTAGGAAGTCAAGAAACACCTGGAGTAA 150754
QY 1010 CAGCAAAATTTGGCTTGGATACGGAATGAGCAGGCAAGGAACTAATAGAGTTTGGC 1069
Db 150753 CAGCAAAATTTGGCTTGGATACGGAATGAGCAGGCAAGGAACTAATAGAGTTTGGC 150694
QY 1070 AAGAAATGCACTGGTATAGCAAAACCCCTTTTCCAAACCAAGAGAGAGTCTTACA 1129
Db 150693 AAGAAATGCACTGGTATAGCAAAACCCCTTTTCCAAACCAAGAGAGTCTTACA 150634
QY 1130 CATGGACATCACAGAGTGTCAACCGGAATCAGTTGATTTATTTCTTTGGAGCCAAA 1189
Db 150633 CATGGACATCACAGAGTGTCAACCGGAATCAGTTGATTTATTTCTTTGGAGCCAAA 150574
QY 1190 GATGGAGAGCTCTATACAGTCAGCAAAACCAAGCAGGAGCTTACTTGGCTCAGATC 1249
Db 150573 GATGGAGAGCTCTATACAGTCAGCAAAACCAAGCAGGAGCTTACTTGGCTCAGATC 150514
QY 1250 ATGAACCTCTTATTTGCCAAATTCAGACTTAATTTGAAGAGTAGGGAACCACTAGAT 1309
Db 150513 ATGAACCTCTTATTTGCCAAATTCAGACTTAATTTGAAGAGTAGGGAACCACTAGAT 150455

QY 1310 CACTCAGGTAAACACCTAAATCCATCCCTTATGATTATACAGTGAAGTGAAGATAGAT 1369
Db 150454 CACTCAGGTAAACACCTAAATCCATCCCTTATGATTATACAGTGAAGTGAAGATAGAT 150395
QY 1370 TTAAAGGGCCCTAGATCTGATAGACAGAGTACCTTAATGAACACTATGGACAGAGTTTCATGACA 1429
Db 150394 TTAAAGGGCCCTAGATCTGAT---AGAGTGCCTGATGAACACTATGGAAACGAGTTTCATGACA 150339
QY 1430 TTGTACAGAGACAGGGATCGAGACCATCCCATCGAAAGAAATGCAAAAGCAAAAT 1489
Db 150338 TTGTACAGAGACAGGGATCGAGACCATCCCATCGAAAGAAATGCAAAAGCAAAAT 150279
QY 1490 GCTGTCTGGGAGGCGCTTTACAAATAGCTGTGAAAAGAGAGAGTGAAGCAAAAGGAA 1549
Db 150278 GCTGTCTGGGAGGCGCTTTACAAATAGCTGTGAAAAGAGAGAGTGAAGCAAAAGGAG 150219
QY 1550 AAAAGGAAAGATATAAGAGCATCTGAATGAGAGTTCCAAA----- 1588
Db 150218 AAAAGGAAAGATATAAGAGCATCTGAATGAGAGTTCCAAAAGAGATATAAG 150159
QY 1589 ----- 1588
Db 150158 AAGCCTTCTTCTGGCATCAATGCAAAAGAAATAGAGGAAACAAACAGAAATGGAAAGACTA 150099
QY 1589 ----- 1588
Db 150098 GGGATCTCTTCAAGAAATCAGAGATACCAAGAGAAACATTTTCATGCAAGATGGGCTCGA 150039
QY 1589 ----- 1588
Db 150038 TAAAGACAGAAATGGTATGACCTTAAACAGAGCAGAGATATTAAAGAGAGATGGCAAG 149979
QY 1589 ----- 1588
Db 149978 ATACACAGAGAACTGTACACAAAGATCTTCATGCCAGATAATCACGATGGTGTGA 149919
QY 1589 ----- 1588
Db 149918 TCACCTGACCTAGAGCCAGACATCTCTGAAATGTGAAGTCAAGTGGGCTTTAGAAAGCATCA 149859
QY 1589 ----- 1588
Db 149858 CTACAAAACAAAGCTAGTGGAGTGAATTCAGATTGAGCTATTCACAAATCCTGAAG 149799
QY 1589 ----- 1588
Db 149798 ATGATGCTGTGAAAGTGTGCACTCAATATGCCAGAAATTTGAAAACTCAGCAGTGGC 149739
QY 1589 ----- 1588
Db 149738 CAGGACCTGGAAAAAGGTGAGTTTTCATTTCCAAATCCCAAGAAAGGCAATGCCAAAGAGT 149679
QY 1589 ----- 1588
Db 149678 GCTCAAACTACCGTACAAATTGCACTCATCTCACACACAGTAAAGTAATGCTCAAAATTC 149619
QY 1589 ----- 1616
Db 149618 TTCAAGCCAGGCTTCATCAATATGGAACCTGTGAACTTCCAGATGTTCAAGCTGGTTTTA 149559
QY 1617 GAAAGTCAAGAGAAACAGACCAAAATTTGCCAATCTCTGTATCATATGAAAGAAAGCAA 1676
Db 149558 GAAAGGACAGAGAACCAAGAAATCAATTTGCCAATCCGCTGGATCGTGGAAAGCA 149499
QY 1677 GAGAGTTCAGAAAAACATCTATTTCTGCTTTATGACTATGCAAAAGCCTTTGACTGTG 1736
Db 149498 GAGAGTTCAGAAAAACATCTATTTCTGCTTTATGACTATGCAAAAGCCTTTGACTGTG 149439
QY 1737 GGGGTCAACAATAAATCTGTGAAAAATTTCTGAAAGGAGTGGGAATACCAACACCACTGACCT 1796
Db 149438 TAGATCACAATTAATCTGTGAAAAATTTCTGAAAGAGATGGGAATACCAACACCACTGACCT 149379
QY 1797 GACTCTTGAAAAATTTGATGCGAGTCAAGGAGCAACAGTTAGAACTGGACATGGAAACA 1856

Db 149378 GCCTCTTGAGAAATTTGTATGTCAGGTGAGAGCAACAGTTGAACTGGACATGGAACAA 149319
QY 1857 CAGACTGGTTCAGAGAGAGAGAGTATGTCAAGCGCTGTATTTGTTCACCGGCTGTCT 1916
Db 149318 TAGACTGGTTCAGAGAGAGAGAGTATGTCAAGCGCTGTATTTGTTCACCGCTGTAT 149259
QY 1917 TTAACCTTCATGAGAG-ACATCATGAGAAACCGCTGGCTGGAGAGACAGCTGGAA 1975
Db 149258 TTAACCTTATGAGAGTATCATCATGAGAAATGCTGACTGGAGAGACAGCTGGAA 149199
QY 1976 TCAAGATTGCGGAGAGAAATAGCAATAACCTCAGATATGAGATATACACCCCTTATGCG 2035
Db 149198 TCAAGATTGCGGAGAGAAATATCAATCACTCAGATATGAGATATACACCCCTTATGCG 149139
QY 2036 CAGAAAGTGAAGAGAACTTAAAGAGCTTTGTATGAAGTGAAGAGGAGAGTGAAGAAAG 2095
Db 149138 CAGAAAGTGAAGAGAACTTAAAGAGCTTTGTATGAAGTGAAGAGGAGAGTGAAGAAAG 149079
QY 2096 TTGGCTTAAGCTCACATTCAGAAACGAGATCATGCTGCTGCCATCCTTCAT 2155
Db 149078 TTGGCTTAAGCTCACATTCAGAAATGAAGATCATGCTGCTGCCATCCTTCAT 149019
QY 2156 GGGAAATAGATGGGAAACAGTGGAAACAGTGCAGACTTTATTTTGGGGGCTCCAA 2215
Db 149018 GGGAAATAGATGGGAAACAGTGGAAACAGTGCAGACTTTATTTT - CTGGGCTCCAA 148961
QY 2216 AATCACTGCAGATGGTGCAGCCATGAAATTAAGACACTTACTCTCTTGGAGAA 2275
Db 148960 AATCACTGCAGATGGTGCAGCCATGAAATTAAGACACTTACTCTCTTGGAGAA 148901
QY 2276 AGTTA--ACCAACTAGATAGCATATTGAAGAGAGACATTTACCTTGGCAACAAAGCG 2333
Db 148900 AGTTATGACCAACTAGATAGCATATTGAAGAGAGACATTTACTTTCACCAAGGT 148841
QY 2334 CCATCTAGTCAAGCTATGTTTTTCCAGTGTGATGTATGATGTGAGATTTGACTGT 2393
Db 148840 TCGTCTAGTCAAGCTATGTTTTTCTGTGGTCAATGATGATGATGAGATTTGACTGT 148781
QY 2394 GAAGAAAGCTGAGCACTGAGATTTGATGCTTTTGAAGTGTGTTGAGAGACTCTT 2453
Db 148780 GAAGAAAGCTGAGCACTGAGATTTGATGCTTTTGAAGTGTGTTGAGAGACTCTT 148721
QY 2454 GAGAGTCCCTTGGACTGCAAGGAGATCCCAACAGTCCATCTCTGAAGAGATCAGCCCTGG 2513
Db 148720 GAGAGTCCCTTGGACTGCAAGGAGATCCCAACAGTCCATCTCTGAAGAGATCAGCCCTGG 148661
QY 2514 GATTTCTTGAAGAGATGATGTAAGCTGAACTCCAGTACTTTGGCCACCTGATCAG 2573
Db 148660 GATTTCTTGAAGAGATGATGTAAGCTGAACTCCAGTACTTTGGCCACCTGATCAG 148601
QY 2574 AAGAGCTGACTCACTGGAAGAGACCTGATGCTGGAGAGGATTTGGGGCAGGAGGAGAG 2633
Db 148600 AAGAGTGTGACTCATTTGAAGAGACTCTGTGCTGGAGGATTTGGGGCAGGAGGAGAG 148541
QY 2634 GGGACGACAGAGGATGATGGCTGGATGTCATCACTGATCGATGAGTGTGCTGGG 2693
Db 148540 GGGACGACAGAGGATGATGGCTGGATGTCATCACTGATCGATGAGTGTGCTGGG 148481
QY 2694 TGAATCCTGAGTGTGATGAGACAGGAGGCTGCTGCTGGCGGATTCATGCGGTTCAC 2753
Db 148480 TGAATCCTGAGTGTGATGAGACAGGAGGCTGCTGCTGGCGGATTCATGCGGTTCAC 148421
QY 2754 AAAGAGTTGACACGACTGAGCAACTGAACTGAACTGAACTG 2795
Db 148420 AAAGAGTTGACAC- - -GAGCGACTGATCTTATCTGATCTG 148383

RESULT 14
AC142564/c 145088 bp DNA linear HTG 02-MAY-2003
LOCUS Bos taurus clone RP42-56B14, WORKING DRAFT SEQUENCE, 2 ordered
DEFINITION pieces.

AC142564
AC142564.2 GI:30315712
HTG: HTGS_PHASE2; HTGS_DRAFT.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 145088)
Antonielli,A., Ayala,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Cariaga,K.,
Chu,G., Coleman,B., Coleman,H., Engle,J., Granite,S., Guan,X.,
Gupta,J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P.,
Huxle,B., Idol,J.R., Karlins,B., Kwong,P., Lalic,P., Lee-Lin,S.-Q.,
Legaspi,R., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Masiello,C.,
Maskeri,B., McDowell,J., Pagnuirgan,C., Pearson,R., Portnoy,M.E.,
Prasad,A., Reddi-Dugue,N., Schandler,K., Schueler,M.G., Shah,K.,
Sison,C., Stanthrop,S., Thomas,J.W., Thomas,P.J., Tsipouri,V.,
Vodt,J.L., Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 145088)
Green,B.D.
Direct Submission
Submitted (04-APR-2003) NIH Intramural Sequencing Center, 8717
Government Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 145088)
Green,E.D.
Direct Submission
Submitted (02-MAY-2003) NIH Intramural Sequencing Center, 8717
Government Circle, Gaithersburg, MD 20877, USA
On May 2, 2003 this sequence version replaced gi:29540599.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoo@nhgri.nih.gov
----- Project Information
Center project name: djr
Center clone name: 056B14

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 144975 bases at least Q40
Consensus quality: 144984 bases at least Q30
Consensus quality: 144986 bases at least Q20
Insert size: 159000; agarose-fp
Insert coverage: 14.02x in Q20 bases; agarose-fp
Quality coverage: 15.37x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 32951: contig of 32951 bp in length
 * 32952 33051: gap of unknown length
 * 33052 145088: contig of 112037 bp in length.

FEATURES

Location/Qualifiers

1..145088

/organism="Bos taurus"

/mol_type="genomic DNA"

/db_xref="taxon:9913"

/clone="RP42-56B14"

/clone_lib="RP42"

1..32951

/note="assembly_fragment"

clone_end:17

vector_side:left"

33052..145088

/note="assembly_fragment"

clone_end:SP6

vector_side:right"

51600..145088

/note="clone overlaps with GenBank Accession Number
 AC141853 clone RP42-374H3 (center project name djp)"

Query Match 58.9%; Score 2103.2; DB 1; Length 145088;

Best Local Similarity 76.9%; Pred. No. 0;

Matches 2613; Conservative 0; Mismatches 168; Indels 518; Gaps 10;

Qy 1 GTCAGGAAGGCGCGAGTGAGGAGAGTACCTTCGTCGAAGGTAAGGAGCAGTAGC 60

Db 9242 GTCAGGAAGGCGCGTGTGAGGAGA-----TACCCTTCGTCGAAGGTAAGGAGCAATGGC 9188

Qy 61 TGCCTTTGCTGGAGCGCGTAAGAGATACCCACGCCCGCCAGGTAAGAGAACCCAG 120

Db 9187 TGCCTTTGCTGGAGCGAGCTGTCAAGAGATACCCACGCTCAAGGTAAGAGAACCCAG 9128

Qy 121 TAAGATGGTAGGTGTTGTGAGAGGGCATCAGAGGGCAGACATCTGAACCATACACGCA 180

Db 9127 TAAGATGGTAGGTGTTCAAGAGCGCATCAGAGGGCAACACACTGAACCATACAC 9068

Qy 181 GAAGACTAGTCATCTAATCACACTAGGACACAGCCTTGTCTAATCAATGAACCTAAG 240

Db 9067 GAAGACTAGTCATCTAATCACACTAGGACACAGCCTTGTCTAATCAATGAACCTAAG 9008

Qy 241 CCATGCCCGTGGGCAACCCCAAGATGGCGAGGTCTCATGGAGAGATCTGACAGATGTG 300

Db 9007 CCATGCCCGTGGGCAACCCCAAGATGGCGAGGTCTCATGGAGAGATCTGACAGATGTG 8948

Qy 301 GTCCACTGGAGAGGGAAAT-GCAACCACTTCAGTATCTTTGCTTCGAGAACCCCATGAA 359

Db 8947 GTCCACTGGAGAGGGAAATGCCAACCACTTCAGTATCTTTGCTTCGAGAACCCCATGAA 8888

Qy 360 CAGTATGAAAGGCAAAATGATAGTACTGAAAGAGGAACTCCCGAGGTTCAGTAGTGC 419

Db 8887 CAGTATGAAAGGCAAAATGATAGTACTGAAAGAGGAACTCCCGAGGTTCAGTAGTGC 8828

Qy 420 CCCATATGCTCAGGATCAGTGGAGAAATAACTCCAGAAAGAAATCAAGAGATGGAGCC 479

Db 8827 CCAATATGCTCAGGATCAGTGGAGAAATAACTCCAGAAAGAAATCAAGAGATGGAGCC 8768

Qy 480 AAAGCAAAAGATACCCGCTGGATGTGATGCTGGTGTATAGCAAGGTCGATGCTG 539

Db 8767 AAAGCAAAAGATACCCGCTGGATGTGATGCTGGTGTATAGCAAGGTCGATGCTG 8708

Qy 540 TAAAGCAATATTGCTAGGAACCTGGAATGTGAGTCCATGAATCAAGGCAAAATTGGA 599

Db 8707 CAAAGCAATATTGCTAGGAACCTGGAATGTGAGTCCATGAATCAAGGCAAAATTGGA 8648

Qy 600 AGTGGTCAAAAGAGATGGCAAGATGATGCAATCTAGGAAATCAGGAACTAAA 659

Db 8647 AGTGGTCAAAAGAGATGGCAAGATGATGCAATCTAGGAAATCAGGAACTAAA 8588

Qy 660 ATGGACTGGGAATGGGTGAAATTAATCAGATGACCATTAATCTACTTGGCGGAGGA 719

Db 8587 ATGGACTGGGAATGGGTGAAATTTAACTCAGATAAACCAATTATATCTACTACTGTGGCAGGA 8528

Qy 720 ATCCCTTCAGAGAAATGGAGTAGCCATCATGTCAACAAAAGAGTCCGAAATGCAGTAGTACT 779

Db 8527 ATGCTTCAGAGAAATGGAGTAGCCATCATGTCAACAAAAGAGTCTGAAATGCAGTAGTACT 8468

Qy 780 TGGATGCAAGTCTCAAAAAACGACAGAAATGATCTCTGTTTTCCTTCCAAAGCAAAACCAATCAA 839

Db 8467 TGGATGCAAGTCTCAAAAAACGACAGAAATGATCTCTGTTTTCCTTCCAAAGCAAAACCAATCAA 8408

Qy 840 TATCAGAGTAATCCCAAGTCTATGCCCCAACAGTAATCTCTGAAGAGCTGAAGTTGACG 899

Db 8407 TATCAGAGTAATCCCAAGTCTATGCCCCAACAGTAATCTCTGAAGAGCTGAAGTTGACG 8348

Qy 900 GTCTATGAAGACCTTACAAGACCTTTTATAGAACTAACACCCCAAAAAGATGCTCTTCTCAT 959

Db 8347 GTCTATGAAGACCTTACAAGACCTTTTATAGAACTAACACCCCAAAAAGATGCTCTTCTCAT 8268

Qy 960 TATAGGAGCTGGAATGCAAAAGTAGGAAGCAAGAAACACCTGGAGTAACAGGCAAAATT 1019

Db 8287 TATAGGAGCTGGAATGCAAAAGTAGGAAGCAAGAAACACCTGGAGTAACAGGCAAAATT 8229

Qy 1020 TGGCCTTCGGAATACGGAATGAAGCAGGCAAGAACTAATAGAGTTTTCGCCAAGAAATGC 1079

Db 8228 TGGCCTTCGGAATACGGAATGAAGCAGGCAAGAACTAATAGAGTTTTCGCCAAGAAATGC 8169

Qy 1080 ACTGTCTATAGCAAAACACCTCTTCCAAACAACAAGAGAGACTCTACATGGACATC 1139

Db 8168 ACTGTCTATAGCAAAACACCTCTTCTAACAACAAGAGAGACTCTATACATGGACATC 8109

Qy 1140 ACCAGATGCTCAACACCGAATCAGATTGATTATATCTTTTCGAGCCAAAGATGAGAGAG 1199

Db 8108 ACCAGATGCTCAACACCGAATCAGATTGATTATATCTTTTCGAGCCAAAGATGAGAGAG 8049

Qy 1200 CTCTATACAGTCAGCAAAACCAAGACCCAGGAGCTTACTGTGGCTCAGATCATGAACCTCT 1259

Db 8048 CTCTATACAGTCAGCAAAACCAAGACCCAGGAGCTTACTGTGGCTCAGATCATGAACCTCT 7989

Qy 1260 TATTGCCAAATCAGACTTAATTTGAAGAAAGTAGGAAACCACTAGATCAGTACAGTA 1319

Db 7988 TATTGCCAAATCAGACTTAAAGTGAAGAAAGTAGGAAACCACTAGATCAGTACAGTA 7999

Qy 1320 AGACCTAAATCCATCCCTTATGATTATACAGTGGAAAGTGAAGAAATAGATTAAAGGCTCT 1379

Db 7928 TGACCTAAATCCATCCCTTATGATTATACAGTGGAAAGTGAAGAAATAGATTAAAGGCTCT 7869

Qy 1380 AGATCTGATAGCAGAGTACCTTATGAACTATGACAGAGGTTTCAATGACATTGTACAGGA 1439

Db 7868 AGATCTGATAGCAGAGTACCTTATGAACTATGAACTATGAGGTTTCTGACATTGTACAGGA 7809

Qy 1440 GACAGGATCAGAGACCTCCCATGGAAAGAAATGCAAAAAGCAAAATGGCTGTCTGG 1499

Db 7808 GACAGGATCAGAGACCTCCCATGGAAAGAAATGCAAAAAGCAAAATGGCTGTCTGG 7749

Qy 1500 GGAGCCTTCAAAATAGCTGTGAAAGAGAGAGTGAAGAAAGCAAAAGGAAAG 1559

Db 7748 GGAGCCTTCAAAATAGCTGTGAAAGAGAGAGTGAAGAAAGCAAAAGGAAAG 7689

Qy 1560 ATAAAGCATCTGAATCAGAGTTCCAAA----- 1588

Db 7688 ATAAAGCATCTGAATCAGAGTTCCAAA----- 7629

Qy 1589 ----- 1588

Db 7628 TAGCGATCAAGGCAAAAAATAGAGGAAAAACAACAGAAATGGAAAGACTAGGATCTCTT 7569

Qy 1589 ----- 1588

Db 7568 CAAGAAATTAGAGATACCAAGGCAATTTCTATGCAAAAGATGGGCTCGATTAAGGACAG 7509

Qy 1589 ----- 1588

Db 7508 AAATGGTATGGACCTTAACAGAGCAGAGATATTAAAGAGAGATGGCAAGATACACAGA 7449

Qy	1589	-----	1588
Db	7448	AGAACTGTATGAAAAAGATCTTCAGACCCAGATAATCACGATGGTGTGATCACTCACCT	7389
Qy	1589	-----	1588
Db	7388	AGAGCCAGACATCCTGGAAATGTGAAGTCAAGTGGCCCTTAGAAGACATCACTACGAACAA	7329
Qy	1589	-----	1588
Db	7328	AGCTAGTGGAGGTGATGAAATTCAGTTGAGCTATTCCAAATCCTGAAGATGATGCTGT	7269
Qy	1589	-----	1588
Db	7268	GAAAGTCTGCACCTCAATATGCCAGCAAAATTTGAAAACCTCAGCAGTGGCCATAGGACTG	7209
Qy	1589	-----	1588
Db	7208	GAAAAGGTCAGTTTTTCATTCCAAATCCCAAGAAAGACAATGCCAAAGAAATGCTCAAACTA	7149
Qy	1589	-----	1588
Db	7148	CCACACAATTGCACCTCATCTCACACACTAGTAAGTAATGCTCAAGTTCTCCAAGCCAG	7089
Qy	1589	-----	1588
Db	7088	GCTTCAGCAATATGTGAACCGTGAACCTTCCTGATGTTCAAGCTGGTTTTAGAAAAGCCAG	7029
Qy	1627	AGGAACAGAGACCAAAATTCGCAACATCCCTCTGTATCATGSAAAAAGCAAGAGAGTTCCA	1686
Db	7028	AGGNACAGAGATCAAAATTCGCAACATCCGCTGGATCATGAAAAGCAAGAGAGTTCCA	6969
Qy	1687	GAAAAACATCTATTTCTGCTTTATTGACTATGCAAAAGCCCTTTGACTGTGGGGGTCACAA	1746
Db	6968	GAAAAGCATCTATTTCTGCTTTATTGACTATGCCAAAGCCCTTTGACTGTGTGGATCACAA	6909
Qy	1747	TAAACTCTGGAATAATCTGAAGGGATGGGAATACAGACCACTGACCTGACTCTTTGAA	1806
Db	6908	GAAAC--GGAAATAATCTGAAGAGATGGGAATACAGACCACTGATCTGCGCTCTTGAG	6851
Qy	1807	AAATTTGTATGCAGGTGAGGAACACAGTTAGAACTGGACATGGAAACACAGACTGGTT	1866
Db	6850	AAATTTGTATGCAGGTGAGGAACACAGTTAGAACTGGACATGGAAACACAGACTGGTT	6791
Qy	1867	CCAGTAGGAAAAGGAGTATGTCGAAGCTCTATATTGTCACCCGGCTTGTTTAACTCTTA	1926
Db	6790	CCAAATAGGAAAAGGAGTTCGTCAAGCTGTGATTTGTTACCTGTTTATTTAACTTATA	6731
Qy	1927	TGCAGAG-ACATCATGAAAAACGCTGGGCTGGGAAGACACAAGCTGGAAATCAAGATTGC	1985
Db	6730	TGCAGAGTACATCATGAAAAATCTGACTGGGAAGAAACAAGCTGGAAATCAAGATTGC	6671
Qy	1986	CGGAGAAATAGCAATTAACCTCAGATATGAGATGATACCACTTATGCGCAAAAGTGA	2045
Db	6670	CGGAGAAAATATCAATAACCTCAGATATGAGATGAGACCA-CTTATGGCAAAAAGTGA	6612
Qy	2046	AGGGAACTAAAAAGCCCTTTGATGAAGGTGAAGAGGAGTGA AAAAGTTGGCTTAAA	2105
Db	6611	AGAGAACTCAAAGCCCTTTGATGAAGTGAAGTGGAGAGTGA AAAAGTTGGCTTAAA	6552
Qy	2106	GCTCAACATTCAGAAAACGAAGATCATGGGATCTGGTCCCATCACTTCATGGGAAATAGA	2165
Db	6551	GCTCAACATTCAGAAAACGAAGATCATGGGATCCGGTCCCATCACTTCATGGGCAAAATAGA	6492
Qy	2166	TGGGAAAACAGTGGAAAACAGTGCAGACTTTATTTTTTGGGGGCTCCAAAATCACTGCA	2225
Db	6491	TGGGAAAACAGTGTCA- -----ACTTTATCTTGGGGGCTCCAAAATCACTACA	6443
Qy	2226	GATGGTGACTGCAGCCATGAAATTAAGAACACTTACTCCTTGGAGAAAAGT-TAACCA	2284
Db	6442	GATGGTGACTGCAGCCATGAAATTAAGAACACTTACTCCTTGGAGAAAAGT-TATGACAA	6383

Qy	2385	ACCTAGATGACATATTGAAAAGCAGAGACATTAACCTTTGCCAACAAGGCCCACTCTAGTCA	2344
Db	6382	ACCTAGATGATGATTCAAAGCAGAGACATTAACCTTTGCCAACAAGGTCCTGCTAGTCA	6323
Qy	2345	AGGCTATGGTTTTTCCAGTGGTTCATGATGATGATGAGCTTGACATGTCGAAGAAGCTG	2404
Db	6322	AGGCTATGGTTTTTCTGTGCTCATGATGATGATGATGAGCTTGACATGTCGAAGAAGCTG	6263
Qy	2405	AGCACTGAAGAATTGATGCTTTTGAACCTGGTGTGTTGGAGAAGACTTTCGAGAGTCCTTT	2464
Db	6262	AGCACCAAGAATTGATGCTTTTGAACCTGGTGTGTTGGAGAAGACTC - GAGAGTCCTTT	6205
Qy	2465	GGACTGCAAGAGATCCAAACAGTCCATCTGAAGAGATCAGCCCTGGATTCCTTTTGG	2524
Db	6204	GGACTGCAAGAGATCCAACTAGTCCATCTGAAGAGATCAGCCCTGGATTCCTTTTGG	6145
Qy	2525	AAGGAATGATGCTAAAGCTGAAATCCAGTACTTTTGGCCACTGATCAGAAGAGCTGACT	2584
Db	6144	AAGGAATGATGCTAAAGCTGAAATCCAGTACTTTTGGCCACTCATGCGAAGAGTTGACT	6085
Qy	2585	CACCTGGAAGACCCCTGATCTCTGGAGGATTCGGGGCAGGAGAGAGGGGACGACAGA	2644
Db	6084	CACCTGGAAGACCTGATCTGATCTGGAGGATTCGGGGCAGGAGAGAGGGGATGACAGA	6025
Qy	2645	GGATGAGATGGCTGGATGGCATCACTGACTCGATGAGCTGAGCTCTGGGTGAACTCCTCG	2704
Db	6024	GGATGAGATGGCTGGATGGCATCACTGACTCGATGAGCTGAGCTGATTTGACTCCGGG	5965
Qy	2705	AGTTGTGATGACAGGAGGCTGCTCTCGCGCATTCATGGGGTCACAAAGATTGGA	2764
Db	5964	TGTTGTGATGACAGGAGGCTCTGCTGATTCATGGGGTTCAAAGAGTCGGA	5905
Qy	2765	CAGGACTGACCACTGAACCTGAACTGAACCTGACTGAAA	2803
Db	5904	CAGGACTGACCACTGATCTGATCTGAGATGTCATCAGA	5866
RESULT 15			
AC142562/c			
LOCUS			
DEFINITION	AC142562	189278 bp DNA linear	HTG 24-MAY-2003
	Bos taurus clone RP42-371H23,	WORKING DRAFT SEQUENCE,	3 unordered
	pieces.		
ACCESSION	AC142562		
VERSION	AC142562.2	GI:31044306	
KEYWORDS	HTG; HTGS PHASE1; HTGS_DRAFT.		
SOURCE	Bos taurus (cow)		
ORGANISM	Bos taurus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelestomii;		
	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;		
	Bovidae; Bos.		
REFERENCE	1 (bases 1 to 189278)		
AUTHORS	Antonellis,A., Ayle,K., Beckstrom-Sternberg,S.M., Benjamin,B.,		
	Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Carliaga,K.,		
	Chu,G., Coleman,B., Coleman,H., Engle,J., Granite,S., Guan,X.,		
	Gupta,J., Haghghi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P.,		
	Hurle,B., Idol,J.R., Karlins,E., Kwong,P., Laric,P., Lee-Lin,S.-Q.,		
	Legaspi,R., Maduro,Q.D., Maduro,V.B., Margulies,E.H., Mastello,C.,		
	Maskari,B., McDowell,J., Paquirigan,C., Pearson,R., Portnoy,M.E.,		
	Prasad,A., Reddix-Dugue,N., Schandler,K., Schueler,M.G., Shah,K.,		
	Sison,C., Stantripop,S., Thomas,J.W., Thomas,P.J., Tsipouri,V.,		
	Vogt,J.L., Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.		
TITLE	NSC Comparative Sequencing Initiative		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 189278)		
AUTHORS	Green,E.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (04-APR-2003) NIH Intramural Sequencing Center, 8717		
	Grovmont Circle, Gaithersburg, MD 20877, USA		
REFERENCE	3 (bases 1 to 189278)		
AUTHORS	Green,E.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-MAY-2003) NIH Intramural Sequencing Center, 8717		
	Grovmont Circle, Gaithersburg, MD 20877, USA		

QY 1917 TTAACCTTCTATGCAGAG-ACATCATGAGAAACGCTGGCTGGAGAGACACAGCTGGAA 1975
 Db 44666 TTAACCTTCTATGCAGAGTACATCATGAGAAATGCTGGACTGGAGAGACACAGCTGGAA 44607
 QY 1976 TCAAGATTGCGGGAGAGAAATAGCAATTAACCTTCAGATATGAGATGATACACACCTTATGG 2035
 Db 44606 TCAAGATTGCGAGAGAAATATCAATTAACCTTCAGATATGAGATGATACACACCTTATGG 44547
 QY 2036 CAGAAAGTGAAGAGAACTAAAAGCCTCTTGATGAAGGTTGAAGAGAGAGTGAAGAAAG 2095
 Db 44546 CAGAAAGTGAAGAGAACTAAAAGCCTCTTGATGAAGGTTGAAGAGAGAGTGAAGAAAG 44487
 QY 2096 TTGCTTTAAAGCTCAACATTCAGAAAACGAGATCATGGATCGTCCCATCACTTCAT 2155
 Db 44486 TTGCTTTAAAGCTCAACATTCAGAAAATGAAGATCAGCATCGGTCCCATCACTTCAT 44427
 QY 2156 GGGAAATAGATGGGAAACAGTGGAAACAGTGTGAGACTTTATTTTGGGGGCTCCAA 2215
 Db 44426 GGGAAATAGATGGGAAACAGTGGAAACAGTGTGAGACTTTATTTT-CTGGGCTCCAA 44369
 QY 2216 AATCACTCAGATGGTGTGACTCGACCATGAATTAAGACACTTACTCTTGGAGAAA 2275
 Db 44368 AATCACTCAGATGGTGTGACTCGACCATGAATTAAGACACTTACTCTTGGAGAAA 44309
 QY 2276 AGTTA--ACCAACCTAGATAGATATTAAGAAAGCAGAGACATTAACCTTGCCAAAGACC 2333
 Db 44308 AGTTATGACCAACCTAGATAGATATTAAGAAAGCAGAGACATTAACCTTGCCAAAGGT 44249
 QY 2334 CAATCTAGTCAAGCTATGTTTTCAGTGGTCAATGATGATGATGATGATGATGATGATGAT 2393
 Db 44248 TCGTCTAGTCAAGCTATGTTTTCAGTGGTCAATGATGATGATGATGATGATGATGATGATGAT 44189
 QY 2394 GAAGAACTGAGCACTCAAGAAATGATGCTTTTGAAGTGGTGGTGGAGAGACTCTT 2453
 Db 44188 GAAGAACTGAGCACTCAAGAAATGATGCTTTTGAAGTGGTGGTGGAGAGACTCTT 44129
 QY 2454 GAGAGTCCCTTGGACTCAAGAGAGATCAACAGTCCATTCGAGAGAGATGAGCCCTGG 2513
 Db 44128 GAGAGTCCCTTGGACTCAAGAGAGATCAACAGTCCATTCGAGAGAGATGAGCCCTGG 44069
 QY 2514 GATTTCTTGAAGAAATGATGCTTAAGCTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 2573
 Db 44068 GATTTCTTGAAGAAATGATGCTTAAGCTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 44009
 QY 2574 AAGAGTCACTCACTGGAAGAGACCTGATGCTGGAGGGATGGGGGAGGAGAGAG 2633
 Db 44008 AAGAGTCACTCACTGGAAGAGACCTGATGCTGGAGGGATGGGGGAGGAGAGAG 43949
 QY 2634 GGGACGACAGAGATGAGATGCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2693
 Db 43948 GGGACGACAGAGATGAGATGCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 43889
 QY 2694 TGAATCTCTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2753
 Db 43888 TGAATCTCTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 43829
 QY 2754 AAAGAGTGGACACAGTGAAGCAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2795
 Db 43828 AAAGAGTGGACAC---GAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 43791

RESULT 17
 AC091252 LOCUS 183383 bp DNA linear MAM 06-OCT-2001
 DEFINITION Bos taurus clone RP42-324C24, complete sequence.
 AC091252 ACCESSION
 AC091252.3 GI:15982451
 HTG
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Bovidae; Bovinae; Bos.
 1 (bases 1 to 183383)
 Akter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Ho, S.-L., Idol, J.R., Karlins, E., Latic, P., Lee-Lin, S.-Q.,
 Legaspi, R., Maduro, Q.L., Maduro, V.B., Masello, C., Mastrian, S.D.,
 McCloskey, J.C., McDowell, J., Pearson, R., Prasad, A., Shevchenko, Y.,
 Stantripop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W.,
 Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Zhang, L.-H. and Green, E.D.
 NISC Comparative Sequencing Initiative
 Unpublished
 Green, E.D.
 Direct Submission
 Submitted (11-APR-2001) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 3 (bases 1 to 183383)
 Green, E.D.
 Direct Submission
 Submitted (06-OCT-2001) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 On Oct 6, 2001 this sequence version replaced gi:14192894.
 ----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: http://www.nisc.nih.gov
 Contact: nisc.mouse@nih.gov
 ----- Project Information
 Center project name: axx
 Center clone name: 324C24

This sequence was finished as follows unless otherwise noted:
 all regions were double-stranded, sequenced with an
 alternate chemistry, or covered by high quality data
 (i.e., phred quality >= 30); an attempt was made to resolve
 all sequencing problems, such as compressions and repeats;
 all regions were covered by at least one plasmid subclone
 or more than one M13 subclone; and the assembly was confirmed
 by restriction digest.

CLONE LENGTH: This sequence represents the entire insert of
 this clone unless otherwise noted. If there are overlapping
 clones, the overlaps are noted in the beginning and end of
 the Features section.

FEATURES
 source
 Location/Qualifiers
 1. 183383
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9913"
 /clone="RP42-324C24"
 /clone_lib="RP42"

misc_feature

/note="this sequence is not the entire insert of clone
 RP42-324C24; clone overlaps with GenBank Accession Number
 AC089993 (nucleotides 138950-162859) clone RP42-553M7
 (center project name axx); this annotated segment
 represents overlap with nucleotides 153711-162859 of
 AC089993"
 4151..4214
 /note="single clone coverage"

misc_feature

104784..104828
 /note="single clone coverage"

misc_feature

105167..105288
 /note="single clone coverage"

misc_feature

135741..135747
 /note="single clone coverage"

misc_feature

160866..160900
 /note="single clone coverage"

Query Match 58.3%; Score 2081.09; DB 1; Length 183383;
 Best Local Similarity 75.5%; Pred. No. 0;

		Matches 2669; Conservative 0; Mismatches 244; Indels 623; Gaps 13;	
QY	6	GAAGGGCGGAGTACAGAGAGGTACCTACCTCGTCCAGGTAAAGAGAGAGTACGTGCGC	65
Db	69888	GAAGGGCGGAGTACAGAGAGTACCTACCTCGTCCAGGTAAAGAGAGAGTACGTGCGC	69942
QY	66	TTTGCTGGAGCAGCGTAAAGAGATACCCACGCCCCAAGGTAAAGAGAGAGTAAAG	125
Db	69943	TTTGCTGGAGCAGCGTAAAGAGATACCCACGCCCCAAGGTAAAGAGAGTAAAG	70002
QY	126	TGTAAGTGTGTGAGAGGCGATCAGAGGCGAGACATCTGAAACCATACACGCAAAA	185
Db	70003	CGGTAGGTGTGTAAGAGGCGATCAGAGGCGAGACATCTGAAACCATACACGCAAAA	70061
QY	186	CTAGTCAATCTAATCAACATAGACACACGCTTGTCTAATCAATGAATCAAGCCATG	245
Db	70062	CTAGTCAATCTAATCAACATAGACACACGCTTGTCTAATCAATGAATCAAGCTATG	70121
QY	246	CCCGTGGGGCAACCCAGATGGGCGAGTCAATGTTGGAGAGATCTGACAGATGTGTTCCA	305
Db	70122	CCCATGGGGCAACCCAGATTTGGTGGGTCTATGTTGGAGAGATCTGACAGATGTGTTCCA	70181
QY	306	CTGGAGAGGGAT- GCAAAACCACTTCAGTATCTTGGCTTGAACCCCATGAAACAGTA	364
Db	70182	CTGGAGAGGGATGCGCAACCACTTCAGTATCTTGGCTTGAACCCCATGAAACAGTA	70241
QY	365	TGAAAAGGCAAAATGATAGGATCTGAAGAGGAATCTCCCGAGTCAGTAGGTGCCCAT	424
Db	70242	TGAAAAGGCAAAATGATAGTACTGAAGAGGAATCTCCCGAGTCAGTAGGTGCCCAT	70301
QY	425	ATGCTACTGAGATCAGTGGAGAAATACTCCAGAAAGATGAAGAGATGGAGCCAAAGC	484
Db	70302	ATGCTACTGAGATCAGTGGAGAAATACTCCAGAAAGATGAAGAGATGGAGCCAAAGC	70361
QY	485	AAAAAGAAATCCAGCTGCTGATGCTGATATAGCAAGGTCCGATGCTGTAAG	544
Db	70362	AAAAAGAAATCCAGCTGCTGATGCTGATATAGCAAGGTCCGATGCTGTAAG	70421
QY	545	AGCAATATTCATAGGAACCTGGAATGTCAGGTCCATGAATCAAGGCAAAATGGAAGTG	604
Db	70422	AGCAATATTCATAGGAACCTGGAATGTCAGGTCCATGAATCAAGGCAAAATGGAAGTG	70481
QY	605	TCAAAAGAGATGGCAAGTGAATGTCACATCTAGGAATCAGCGCAACTAAATGGA	664
Db	70482	TCAAAAGAGATGGCAAGTGAATGTCACATCTAGGAATCAGCGCAACTAAATGGA	70541
QY	665	CTGGAATGGGTGAATTTAACTCAGATGACCATATATCTACTGCGGGCAGGAATCCC	724
Db	70542	CTGG-ATGGGTGACTTTAACTCAGATGACCATATATCTACTGCGGGCAGGAATCCC	70600
QY	725	TCAGAGAAATGGAGTAGCCATCATGTCACAAAAGTCCGAAATGCAAGTACTTGGAT	784
Db	70601	TCAGAGAAATGGAGTAGCCATCATGTCACAAAAGTCTGAAATACAGTACTTGGAT	70660
QY	785	GCGTCTCAAAAACGACAGATGATCTCTGTTGTTTCCAAAGGCAACCATTCATATCA	844
Db	70661	GCAATTTCAAAAACGACAGATGATCTCTGTTGTTTCCAAAGGCAACCATTCATATCA	70720
QY	845	CAGTAATCCAAGTCTATGCCCAACCAAGTAAATGCTGAAGAGCTGAAGTTGAACGCTCT	904
Db	70721	CAGTAATCCAAGTCTATGCCCAACCAAGTAAATGCTGAAGAGCTGAAGTTGAACGCTCT	70780
QY	905	ATGAGACCTTACAGACCTTTTAGAATCAACCC-AAAAAAGATGCTCTTCTCATATTA	963
Db	70781	ATGAGACCTTACAGACCTTTTAGAATCAACCCAAAAAAGATGCTCTTCTCATATTA	70840
QY	964	GGGACTTGAATGCMAAAGTAGGAAGCAAAACACCTCGAGTAAACAGGCAAAATTTGGC	1023
Db	70841	GGGACTTGAATGCMAAATAGGAAGTCAAGAAACACCTCGAGTAAAGGCAAAATTTGGC	70900
QY	1024	CTTGGAAATACGGAATGAAGAGGCGCAAGACTAATAGAGTTTGGCAGAAATGCACTG	1083
Db	70901	CTTGGAAATACGGAATGAAGAGGCGCAAGACTAATAGAGTTTGGCAGAAATGCACTG	70960

QY	1084	GTCAATAGCAAAACACCTCTTCCAAACACACAAAGAGAGACTTACACATGACATCACCA	1143
Db	70961	GTCAATAGCAAAACACCTCTTCCAAACACACAAAGAGAGACTTACATGACATCACCA	71020
QY	1144	GATGTGTCAACCCGAAATCAGATTATATCTTTTCAGCCAAAGATGGAGAGCTCT	1203
Db	71021	GATGTGTCAACCCGAAATCTGATTATATCTTTTCAGCCAAAGATGGAGAGCTCT	71080
QY	1204	ATACAGTCAGCAAAAACAGACCCAGGAGCTTACTGTGGCTCAGATCATGAATCTCTTAT	1263
Db	71081	ATACAGTCAGCAAAAACAGACCCAGGAGCTGACTGTGGCTCAGACCATGAATCTCTTAT	71140
QY	1264	GCCAAATTCAGACTTAAATTTGAAGAAAGTAGGAAACCACTAGATCACTCAGGTAAGAC	1323
Db	71141	GCCAAATTCAGACTTAAATTTGAAGAAAGTAGGAAACCACTAGACCATTCAGGTATGAC	71200
QY	1324	CTAAATCCAAATCCCTTATGATTATACAGTGAAGTGAAGAAATAGATTTAAGGGCTAGAT	1383
Db	71201	CTAAATCCAAATCCCTTATGATTATACAGTGAAGTGAAGAAATAGATTTAAGGGCTAGAT	71260
QY	1384	CTGATAGACAGAGTACCTTAATGAACCTATGGACAGAGGTTTCATGACATTTGACAGAGACA	1443
Db	71261	CTGATAGATAGAGTGCCTGATGAATGGAATGAGGTTTCATGACATTTGACAGAGACA	71320
QY	1444	GGGATCGAGACCCATCCCATGGAAAAAATGCAAAAAAGCAAAATGGCTGTCTGGGGAG	1503
Db	71321	GGGATCAAGACCATCCCATGGAAAAAATGCAAAAAAGCGAAATGCTGTCTGGGGAG	71380
QY	1504	GCCTTACAATAGCTGTGAAAAGAGAGAGTGAAGCAAGCAAGCAAGCAAGCAAGCAAG	1563
Db	71381	GCCTTACAATAGCTGTGAAAAGAGAGAGTGAAGCAAGCAAGCAAGCAAGCAAGCAAG	71437
QY	1564	AAGCATCTGAATGACAGAGTTCCAAA-	1588
Db	71438	AAACATCTGAATGACAGAGTTCCAAAAGATAGCAAGAGATAGCAAGAGCTTCTTCAGC	71497
QY	1589	-----	1588
Db	71498	GATCAATGCAAAAGAAATAGAGGAAAAACAAGAAATGGGAAAGACTAGGGAATCTCTTCAAG	71557
QY	1589	-----	1588
Db	71558	AAATCAGAGATACCAAAAGGAACATTTTCATGCAAAAGATGGGCTCAATAAAGACAGAAAT	71617
QY	1589	-----	1588
Db	71618	GGTATGGACCTAACAGAGACAGAGATATTAAGAGACATGCGCAATATACACAGAGAA	71677
QY	1589	-----	1588
Db	71678	CTGTACAGAAAAGATCTTTCATGACCCAGATAATCAATGGTGTGATCACTGACCTAGAG	71737
QY	1589	-----	1588
Db	71738	CCAGACATCTCGAATGTGAAGTCAAGTGGGCTTAGAAGCATCACTACAAACAAAGCT	71797
QY	1589	-----	1588
Db	71798	AGTGAGGTGATGGAATTCAGATTGAGCTATTCCAAATCTCAAAAGATGATCTGTGAAA	71857
QY	1589	-----	1588
Db	71858	GTGTTGCACTCAATATGCCAGACATTTTGGAAAACTCAGCAGTGGCCACAGACTGGAAA	71917
QY	1589	-----	1588
Db	71918	AGGTCAGTTTTTCATTTCCAAATCCCAAAGAAAGCAATGCCAAAGATGCTCAAACTACTGC	71977
QY	1589	-----	1588
Db	71978	ACAATTCAGCTCACTCAGCTAGTAAAGTAAATGCTCAAAATTTCTCCAAAGCCAGGCTT	72037

QY 1589 -----GAACTCCAGTGTGTTCAAGCTGGTTTTAGAAAGTCAGAGG 1630
 Db 72038 CAGCAATATGTGAACCATGAATTTCTTGATGTTCAAGCTGGTTTTAGAAAGTCAGAGG 72097
 QY 1631 ACCAGAGACCAAAATGTGCAACATCCTCTGTATCATGTGAAAAGCAAGAGAGTTCCAGAAA 1690
 Db 72098 ACCAGAGATCAATATGGCCCAACATCCGCTGGATCATGGAAGAAAGCAAGAGAG-TCCAGAAA 72156
 QY 1691 AACATCTATTTCTGCTTTATGTACTATGCAAAAGCCTTTCACTGTGGGGTCTCAATAAA 1750
 Db 72157 AGCATCTATTTCTGCTTTATGTACTATGCAAAAGCCTTTCACTGTGGGGTCTCAAGAAA 72216
 QY 1751 CTGTGAAAATTTCTGAAGGATGGGAATACCAAGACCACTGACCTTCTTGAAGAAAT 1810
 Db 72217 CTGTGAAAATTTCTGAAGATGGGAATACCAAGACCACTGACCTTCTTGAAGAAAT 72276
 QY 1811 TTGTATGAGTTCAGGAAGCAAGTAGTTAGAACTGGAACATGGAACACAGACTGGTTCCAA 1870
 Db 72277 TTGTATGAGTTCAGGAAGCAAGTAGTTAGAACTGGAACATGGAACACAGACTGGTTCCAA 72336
 QY 1871 GTAGAAAAGGAGTATGTCAAGGCTGTATATGTCAACCGGCTTGTTTAACTTTCTATGCA 1930
 Db 72337 ATAGAAAAGGAGTTCATCAAGGCTGTATATGTCAACCGTGTATTTAACTTTATATGCA 72396
 QY 1931 GAG-ACATCATGAGAAACGCTGGCTGGAGAGACACAACTGGAATCAAGATTGCGGG 1989
 Db 72397 GAGTACATCATGAGAAACGCTGGACTGGAGAGAAACACAACTGGAATCAAGATTGCGAGG 72456
 QY 1990 AGAATAGCAATACCTCAGATATGAGATGATACCAACCTTTATGSCAGAAAGTGAGAG 2049
 Db 72457 AGAATATCAATACCTCAGATATGAGATGATACCAACCTTTATGSCAGAAAGTGAGAG 72516
 QY 2050 GAACTAAAAGCCCTCTTGATGAAGCTGAAAGAGAGAGAGTGAAGAGTTGGCTTTAAAGCTC 2109
 Db 72517 GAACTAAAAGCCCTCTTGATGAAGCTGAAAGAGAGAGTGAAGAGTTGGCTTTAAAGCTC 72576
 QY 2110 AACATTCAGAAAACGAGATCATGGCATCTGTGTCCTCATCTTCATGCGGAAATAGATGGG 2169
 Db 72577 AACATTCAGAAAACGAGATCATGACATCCGTCCTCATCTTCATGCGGAAATAGATGG- 72635
 QY 2170 GAAACAGTGAAGACAGTGTGAGCTTTATTTTGGGGGCTCCAAATCACTGCAAGATG 2229
 Db 72636 -----GGAACAGTGTGAGCTTTATTTTCTGGGCTCCAAATCACTACAGATG 72685
 QY 2230 GTGACTGAGCCATGAAATTAAGACACATCTACTCTTGGAGAGAAAGTTA--ACCAACC 2287
 Db 72686 GTGACTTCAGCCATGAAATTAAGACAGCTTACTCTTGGAGAGAAAGTTATGACCAACC 72745
 QY 2288 TAGATGATATTAAGAGAGAGACATTAACCTTGCCCAACAAAGCCCATCTAGTCAAGG 2347
 Db 72746 TAGATGATATTAAGAGAGAGACATTAACCTTGCCCAACAAAGGTCATCTAGTCAAGG 72805
 QY 2348 CTATGCTTTTCCAGTGGTATGATGATGTGAGAGTTGAGCTGTGAGAGAAAGCTGAGC 2407
 Db 72806 CTATGCTTTTCC--TGTCATGATGATGTGAGAGTTGAGCTGTGAGAGAAAGCTGAGT 72862
 QY 2408 ACTGAAGATTTGATGCTTTTGAATGCTGTGAGAGAAAGCTTTGAGAGTCCCTTGA 2467
 Db 72863 GCTAAAGAAATTTGATGCTTTTGAATGCTGTGAGAGAAAGCTTTGAGAGTCCCTTGA 72922
 QY 2468 CTGAAGAGATCAACACAGTCCATTTCTGAAGAGATCAGCCCTGGGATTTCTTTGAAG 2527
 Db 72923 CTGAAGAGATCAACACAGTCCATTTCTGAAGAGATCAGCCCTGGGATTTCTTTGAAG 72982
 QY 2528 GAATGATGCTAAAGCTGAACCTCAGTACTTTGGCCACCTGATCAGAAAGCTGACTCAC 2587
 Db 72983 GAATGATGCTAAAGCTGAACCTCAGTACTTTGGCCACCTCATGCAAAAGAGTTGACTCAT 73042
 QY 2588 TGGAAAAGACCTCATGCTGGAGGGATTTGGGGCAGAGAGAGAGGGGACGACAGAGGA 2647
 Db 73043 TGGAAAAGACCTCATGCTGGAGGGATTTGGGGCAGAGAGAGAGGGGATGACAGAGGA 73102
 QY 2648 TGAGATGGCTGGATGGCATCACTGACTCGATGGAGCTGAGTCTGGGTGAACCTCTCTGGAGT 2707

Db 73103 TGAGATGCCCGATGGCATCACTGACTCAGTGGATGTGAGTCTGGTGAATTCGGGAGT 73162
 QY 2708 TGGTATGAGACAGGAGCGCTGTCTCGGGGATTCATGGGTCACAAAGAGTTGGACAC 2767
 Db 73163 TGGTATGAGACAGGAGCGCTGTCTCGGGGATTCATGGGTCACAAAGAGTTGGACAT 73222
 QY 2768 GACTCAGCAACTGAACTCAACTGAACTGTACTGAAACCTTAGTAGTTTATATTACTCAGA 2827
 Db 73223 GACTCAGGACGATGATCTGATCTGACTGACTACTATTTTCAGACTGTTTGTGACTG 73282
 QY 2828 AAATAGTAATTTCAATATTTTCAAAATTTATTTCAATATTTTGGTTAAGATAAAGATT 2887
 Db 73283 TGAGGGCTACTCCATTTCTTCAAAAGGCATTTCTGCCACAGCTAGTAGATATAATGGTCA 73342
 QY 2888 TTCAAAATTTGATTTTATCTTTGATTTTCTCTACTTATTTTATTTTGGATTTAA 2943
 Db 73343 CTGATTTAAATTTCCCATTCAGTCTATTTTGTCTGATTTCCCAAAATTTGA 73398

RESULT 18

AC146889 139617 bp DNA linear HTG 04-NOV-2003
 Bos taurus clone RP42-509J20, WORKING DRAFT SEQUENCE, 4 ordered
 pieces.

ACCESSION

AC146889

VERSION

AC146889.2 GI:38153770

KEYWORDS

HTG; HTGS_PHASE2; HTGS_DRAFT.

SOURCE

Bos taurus (cow)

ORGANISM

Bos taurus

REFERENCE

1 (bases 1 to 139617)

AUTHORS

Antonellis, A., Avele, K., Benjamin, B., Blakesley, R. W.,
 Bouffard, G. G., Brinkley, C., Brooks, S., Chu, G., Coleman, B.,
 Coleman, H., Engle, J., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Han, J., Hansen, N., Ho, S.-L., Hu, P., Hurie, B., Idol, J. R.,
 Karlins, E., Kwong, P., Laric, P., Larson, S., Lee-Lin, S.-Q.,
 Legaspi, R., Maduro, Q. L., Maduro, V. B., Margulies, E. H., Masiello, C.,
 Maskeri, B., McDowell, J., Mullikin, J. C., Paguirigan, C., Pearson, R.,
 Portnoy, M. E., Prasad, A., Reddix-Dugue, N., Schandler, K.,
 Schueler, M. G., Shah, K., Sison, C., Stantripop, S., Thomas, J. W.,
 Thomas, P. J., Tsipouri, V., Vogt, J. L., Wetherby, K. D., Young, A. and
 Green, E. D.

TITLE

NISC Comparative Sequencing Initiative

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 139617)

AUTHORS

Green, E. D.

JOURNAL

Direct Submission

Submitted (16-OCT-2003) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA

3 (bases 1 to 139617)

REFERENCE

Green, E. D.

AUTHORS

Direct Submission

Submitted (04-NOV-2003) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA

On Nov 4, 2003 this sequence version replaced gi:37693553.

COMMENT

----- Genome Center

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: <http://www.nisc.nih.gov>

Contact: nisc.zoo@hgr.nih.gov

----- Project Information

Center project name: eyr

Center clone name: 509J20

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g.,

human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 138458 bases at least Q40
 Consensus quality: 138835 bases at least Q30
 Consensus quality: 139121 bases at least Q20
 Insert size: 125000; agarose-fp
 Insert size: 139317; sum-of-contigs
 Quality coverage: 15.38x in Q20 bases; agarose-fp
 Quality coverage: 13.80x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and the accession number will be preserved.

* 1 59777: contig of 59777 bp in length
 * 59778 114249: gap of unknown length
 * 59778 114249: contig of 54372 bp in length
 * 114350 135352: gap of unknown length
 * 114350 135352: contig of 21003 bp in length
 * 135353 135452: gap of unknown length
 * 135453 139617: contig of 4165 bp in length.

FEATURES

source

1. 139617

/organism="Bos taurus"

/mol_type="genomic DNA"

/db_xref="taxon:9913"

/clone="RP42-509J20"

/clone_lib="RP42"

misc_feature

1. 59777

/note="assembly_fragment"

clone_end:SP6

vector_side:left

59778..114249

/note="assembly_fragment"

114350..135352

/note="assembly_fragment"

135453..139617

/note="assembly_fragment"

clone_end:T7

vector_side:right

Query Match 58.08; Score 2071.6; DB 1; Length 139617;
 Best Local Similarity 76.63; Pred. No. 0;
 Matches 2596; Conservative 0; Mismatches 169; Indels 626; Gaps 11;

Qy	1	GTACGAGGCGGCGCAGTGGAGGAGGTACCTACCTCGTCCCAAGGTAAAGGACGATGC	60
Db	29610	GTACGAGGCGTGGCGGTGAGGAGA-----TACCCCTCGTCCAAAGGTAAAGGACGATGC	29664
Qy	61	TGGCCTTTGTGGAGCGCGGTAAAGAGATACCCAGCCGCGCAAGGTAAAGAAACCCCAAG	120
Db	29665	TGTGCTTTGTGGCAGCGCGGTAAAGAGATACCCCATGCCCAAGGTAAAGAAACCCCAAG	29724
Qy	121	TAAGATGGTGGTGTGTGGAGGGGCATCAGAGGGGAGACATCTGAAACCATACAGCA	180
Db	29725	TAAGATGGTGGTGTGTGGAGGGGCATCAGAGGGGAGACATCTGAAACCATACAGCA	29784
Qy	181	GAAGAACTAGTCAATCTAATCAGTACCTAGGACCAAGCGCTTGTCTAATCTAATGAACTAAG	240
Db	29785	GAAGAACTAGTCAATCTAATCAGTACCTAGGACCAAGCGCTTGTCTAATCTAATGAACTAAG	29844

Qy	241	CCATGCCCGTGGGCAACCCCAAGATGGCAGCTCATGGTGAGAGATCTGACAGATGTG	300
Db	29845	CCATGCCCGTGGGCAACCCCAAGATGGTGGGTCTATGGTGAGAGATCTGACAGATGTG	29904
Qy	301	GTCCACTGGAGAGGGAAT-GCAAAACCACTTCAGTATTCTTTCCTTGAGAAACCCCATGAA	359
Db	29905	GTCCACTGGAGAGGGAATGCAAAACCACTTCAGTATTCTTTCCTTGAGAACTCCATGAA	29964
Qy	360	CAGTATGAAAAGGCAAAATAGTAGTACTGAAAGAGAACTCCCCAGGTCAAGTAGTGC	419
Db	29965	CAGTATGAAAAGGCAAAATAGTAGTACTGAAAGAGAACTCCCCAGGTCAAGTAGTGC	30021
Qy	420	CCCATATCTACTGAGATCAGTGGAGAAATACTCCAGAAAGAAATGAAGAGATGGAGCC	479
Db	30022	-----TGCTACTGGAGATCAGTGGAGAAATACTCCAGAAAGAAATGAAGAGATGGAGCC	30075
Qy	480	AAAGCAAAAGAATAACCCAGCTGGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	539
Db	30076	AGGGCAAAACAATACCCAGCTGGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	30135
Qy	540	TAAAGAGCAATATTGCATAGGAACCTGGAATGTCAGGTCCATGATCAAGCAAAATGGA	599
Db	30136	TAAAGAGCAATATTGCATAGGAACCTGGAATGTCAGGTCCATGATCAAGCAAAATGGA	30195
Qy	600	AGTGTCTCAACAAGAGATGGCAAGATGCAATGTCAACATTTCTAGGAATCAGCGAACTAAA	659
Db	30196	AGTGTCTCAACAAGAGATGGCAAGATGCAATGTCAACATTTCTAGGAATCAGCGAACTAAA	30255
Qy	660	ATGACTCGAATGGGTGAATTTAATCAGATGACCATATATATCTACTACTGCGGCGAGGA	719
Db	30256	ATGACTCGAATGGGTGAATTTAATCAGATGACCATATATATCTACTACTACTACTACT	30313
Qy	720	ATCCCTCAGAGAAATGGAGTAGCCATCATGCTCAACAAAAGAGTCCGAAATGCACTACT	779
Db	30314	ATCCCTCAGAGAAATGGAGTAGCCATCATGCTCAACAAAAGAGTCCGAAATGCACTACT	30373
Qy	780	TGGATGCACTCTCAAAAACGACAGAAATCATCTGTTTGTGTTTCCAAAGCAAAACCAATCAA	839
Db	30374	TGGATGCACTCTCAAAAACGACAGAAATCATCTGTTTGTGTTTCCAAAGCAAAACCAATCAA	30433
Qy	840	TATCAGTATCTCAAGTCTATGCCCCCAACCAAGTAATGCTGAAGAGCTGAAAGTGAAGC	899
Db	30434	TATCAGTATCTCAAGTCTATGCCCCCAACCAAGTAATGCTGAAGAGCTGAAAGTGAAGC	30493
Qy	900	GTCCCTATGAAGACCTTACAAAGACCTTTTAAAGAACTAACACCCCAAAAAGATGCTCTCAT	959
Db	30494	GTCTCTATGAACATCTACAAAGACCTTTTAAAGAACTAACACCCCAAAAAGATGCTCTCAT	30553
Qy	960	TATAGGGACTTGAATGCAAAAGTAGGAAGCAAGAAACACCTGGAGTAAAGGCAAAAT	1019
Db	30554	TATAGGGACTTGAATG-AAAAGTAGGATGTCAAGAAACACCTGGAGTAAAGGCAAAAT	30612
Qy	1020	TGGCCTTGGATACGGAATGAAGCAGGGCAAGAACTAATAGAGTTTTCCTCAAGAAATGTC	1079
Db	30613	TGGCCTTGGATATGGAATGAAGCAGGGCAAGAACTAATAGAGTTTTCCTCAAGAAATGTC	30672
Qy	1080	ACTGTCTATAGCAAAACACCTCTTCCAAACAACAAGAGAGACTCTACATGAGACATC	1139
Db	30673	ACTGTCTATAGCAAAACACCTCTTCCAAACAACAAGAGAGACTCTACATGAGACATC	30732
Qy	1140	ACCAGATGCTCAACACCCCAAAATCAGATTGATTATTTCTTTGACGCCAAGAGTGGAGAG	1199
Db	30733	ACCAGATGCTCAACACTTAAATCAGATTGATTATTTCTTTGACGCCAAGAGATGGAGAA	30792
Qy	1200	CTCTATACAGTCAGCAAAACCAAGACCCAGGAGCTTACTGTGGCTCAGATCATGAACTCCT	1259
Db	30793	CTCTATACAGTCACAAACCAAGACCCAGGAGCTTACTGTGGCTCAGATCATGAACTCCT	30852
Qy	1260	TATTCGCAAAATTCAGACTTAAATGAAGAGTAGGGGAAACCACTAGATCACTCAGGTA	1319
Db	30853	TATTCGCAAAATTCAGACTTAAATGAAGAGTAGGGGAAACCACTAGATCACTCAGGTA	30912

AC145440
AC145440.2 GI:33413342
HTG: HTGS_PHASE2; HTGS_DRAFT.
Bos taurus (ccw)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 172022)
Antonnellis,A., Ayela,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Cariaga,K.,
Chu,G., Coleman,B., Coleman,H., Engle,J., Granite,S., Guan,X.,
Gupta,J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P.,
Hurle,B., Idol,J.R., Karlins,E., Kwong,P., Laric,P., Lee-Lin,S.-Q.,
Legaspi,K., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Masello,C.,
Mascari,B., McDowell,J., Pagnuirgan,C., Pearson,R., Portnoy,M.E.,
Prasad,A., Reddi-Dugue,N., Schardier,K., Schueler,M.G., Shah,K.,
Sison,C., Stantirip,S., Thomas,J.W., Thomas,P.J., Tsipouri,V.,
Vogt,J.L., Wetherby,K.D., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 172022)
Green,E.D.
Direct Submission
Submitted (09-JUL-2003) NIH Intramural Sequencing Center, 8717
Government Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 172022)
Green,E.D.
Direct Submission
Submitted (02-AUG-2003) NIH Intramural Sequencing Center, 8717
Government Circle, Gaithersburg, MD 20877, USA
On Aug 2, 2003 this sequence version replaced gi:32480636.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: <http://www.nisc.nih.gov>
Contact: nisc_zoonhgri.nih.gov
----- Project Information
Center project name: edx
Center clone name: 116J02

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

Sequencing summary statistics
 Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 171033 bases at least Q40
 Consensus quality: 171360 bases at least Q30
 Consensus quality: 171567 bases at least Q20
 Insert size: 184000; agarose-
 Insert size: 171622; sum-of-
 Quality coverage: 11.81x in Q20 bases; agarose-
 Quality coverage: 11.28x in Q20 bases; sum-of-
 Quality coverage: 11.28x in Q20 bases; sum-of-
 Quality coverage: 11.28x in Q20 bases; sum-of-

```

* ----- This is a 'working draft' sequence. It currently
* consists of 5 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and

```

- * This sequence will be replaced
- * by the finished sequence as soon as it is available and

* the accession number will be preserved.	
	1 61497: contig of 61497 bp in length
	61498 61597: gap of unknown length
	61598 100735: contig of 39138 bp in length
	100736 100835: gap of unknown length
	100836 114148: contig of 13313 bp in length
	114149 114248: gap of unknown length
	114249 147871: contig of 33623 bp in length
	147872 147971: gap of unknown length
	147972 172022: contig of 24051 bp in length.
FEATURES	Location/Qualifiers
source	1..172022
	/organism="Bos taurus"
	/mol_type="genomic DNA"
	/db_xref="taxon:9913"
	/clone="RP42-116J2"
	/clone_lib="RP42"
misc_feature	1..61437
	/note="assembly_fragment"
	clone_end:T7
	vector_side:left
misc_feature	1..49037
	/note="clone overlaps with GenBank Accession Number AC140094 clone RP42-369C3 (center project name edy) "
misc_feature	61598..100735
	/note="assembly_fragment"
misc_feature	100836..114148
	/note="assembly_fragment"
misc_feature	114249..147871
	/note="assembly_fragment"
misc_feature	147972..172022
	/note="assembly_fragment"
	clone_end:SP6
	vector_side:right
Query Match	58.0%; Score 2071.2; DB 1; Length 172022;
Best Local Similarity	77.0%; Pred. No. 0;
Matches 2524; Conservative	0; Mismatches 148; Indels 608; Gaps 6;
Qy	112 AAACCCAGTAAGATGTAGTGTGTTCAGAGGGCATCAGAGGGCGAGACATCTGAAACC 171
Db	147703 AGAGACAGTAAGATGTAGTGTGTTCAGAGGGCATCAGAGGGCGAGACATCTGAAACC 147644
Qy	172 ATACACGCAGAAAACTAGTCAATCTAATCATCACATAGGACCACAGCCTTGTCCTCAAT 231
Db	147643 ATACTCAGAAAACTAGTCAATTTAATCATCACATAGGACTACAGCCTTCTCTAACTCAAT 147584
Qy	232 GAAACTTAGCCATGCCGTGGGGCAACCCAGATGGGCAGCTCATGTGGAGAGATCTGA 291
Db	147583 GAACTAGCCATGCCCATAGAGAACCCAGATGGCGGTCTATGTGGAGATCTGA 147524
Qy	292 CAGAATGTGGTCCACTGGAGAAGGAAAT-GCAACCACTTCAGTATTCCTTGCCCTTGAGAA 350
Db	147523 CAGAATGTGGTCCACTGCAGAAGGAAATGGCAACCACTTCAGTATTCCTTGCCCTTGAGAA 147464
Qy	351 CCCCATGACAGTATGAAAGGCAAAATGATAGATACTCAAGAGGAACTCCCGAGTTC 410
Db	147463 CCCCATGACAGTATGAAAGGCAAAATGATAGATACTCAAGAGGAACTCCCGAGTTC 147404
Qy	411 AGTAGTGCCCCATATGCTACTCTGGAGATCAGTGAGAAATAAATCCAGAAAGAAATGAAGA 470
Db	147403 AGTAGTGCCCCAATATGCTACTCTGGAGATCAGTGAGAAATAAATCCAGAAAGAAATGAAG 147384
Qy	471 GATGAGCCAAAGCAAAAAGAAATACCCAGCTGTGGATGTGACTGTGTGATATAGCAAGGT 530
Db	147343 GATGGAGCAAAAGCAAAAAGCAATATCCAGTGTGTGGATGTGACTGTGGTGAAGCAAGGT 147284
Qy	531 CGGATCTCTAAAGAGCAATATTCATAGGAACCTGGAATGTTCAGGTCCCATGATCAAGG 590
Db	147283 CGGATCTCTAAAGAGCAATATTCCTTAGAACCTGGAATGTTCAGGTCCCATGATCAAGG 147224
Qy	591 CAAATTTGGAAGTGGTCAAAACAGAGATGGCAAGTAGTGAATGTCAACATTTAGGAATTCAG 650

Db	147223	CAAAATGGAAGTGGTCAAAACAGAGATGGCAAGAGTGAATGTTGACATTTCTAGGAATCGG	147164
QY	651	CGAACTAAATGGAGTGGATGGGTGAATTTAACTCAGATGACCATTTATCTACTACTG	710
Db	147163	CAAACTAAATGGAGTGGATGGGTGAATTTAACTCAGATGACCATTTATCTACTACTG	147104
QY	711	CGGCGAGGAATCCCTCAGAGAAATGGAGTAGCCATCATGGTCAACAAAAGAGTCCGAAA	770
Db	147103	TGGCAGGAATCCCTCAGAGAAATGGAGTAGCCATCATGGTCAACAAAAGAGTCTCAAG	147044
QY	771	TGCAGTACTTGGATGCGAGTCTCAAAAACGACAGAAATGATCTCTGTTTTCMAAGCAA	830
Db	147043	TGCAGTACTTGGATGCGAGTCTCAAAAACGACAGAAATGATCTCTGTTTTCMAAGCAA	146984
QY	831	ACCAATCAATACAGTAATCCAAAGTCTATGCCCCCAACAGTAATGCTGAAGAAGCTGA	890
Db	146983	ACCAATCAATACAGTAATCCAAAGTCTATGCCCCCAACAGTAATGCTGAAGAAGCTGA	146924
QY	891	AGTTGAAGCGTCTATGAAGACCTTACAGACCTTTTGAAGCTTAACACCCAAAAGATGT	950
Db	146923	AGTTGAAGCGTCTATGAAGACCTTACAGACCTTTTGAAGCTTAACACCCAAAAGATGT	146864
QY	951	CTTTCTCATATAGGGGACCTGGATGCAAAAGTAGGAAGCAAGACACCTCGAGTAAC	1010
Db	146863	CTTTCTCATATAGGGGACCTGGATGCAAAAGTAGGAAGCTCAAGAACACCTCGAGTAAC	146804
QY	1011	AGGCAAAATTTGGCCCTTGGAAATACGGAAATGAAGCAGGGCAAGACTAATAGAGTTTGGCA	1070
Db	146803	AGGCAAAATTTGGCCCTTGGAAATACGGAAATGAAGCAGGGCAAGACTAATAGAGTTTGGCA	146744
QY	1071	AGAAATGCACTGCTATAGCAACACCTCTTCCACACACACAGAGAGAGCTCTACAC	1130
Db	146743	AGAAATGCACTGCTATAGCAACACCTCTTCCACACACACAGAGAGAGCTCTACAC	146684
QY	1131	ATGACACATCACAGATGGTCAACACCGAAATCAGATTGATTATATCTTTGCCAGCCAAAG	1190
Db	146683	ATGACACATCACAGATGGTCAACACCGAAATCAGACTGATTATATCTTTGCCAGCCAAAG	146624
QY	1191	ATGGAGAGCTCTATACAGTCAGCAAAAACAGACCGAGGAGCTTACTGTGGCTCAGATCA	1250
Db	146623	ATGGAGAGCTCTATACAGTCAGCAAAAACAGACCGAGGAGCTTACTGTGGCTCAGATCA	146564
QY	1251	TGAATCTCTTATTCGCAAAATTCAGACTTAAATTCAGCAAAAGTAGGGGAAAACCACTAGATC	1310
Db	146563	TGAATCTCTTATTCGCAAAATTCAGACTTAAATTCAGCAAAAGTAGGGGAAAACCACTAGATC	146504
QY	1311	ACTCAGTAAGACCTAAATCCATCCCTTATGATTATACAGTGGAGTGAAGAAATAGATT	1370
Db	146503	ACTCAGTAAGACCTAAATCCATCCCTTATGATTATACAGTGGAGTGAAGAAATAGATT	146444
QY	1371	TAAAGGCTAGATCTAGTACAGTACCTTAATGAACTATGGACAGAGCTTCTATGACAT	1430
Db	146443	TAAAGGCTAGATCTAGTACAGTACCTTAATGAACTATGGAACTATGGATGAGTTCTATGACAT	146384
QY	1431	TGTACAGAGACAGGGATCGAGACCATCCCATGGAAAAGAAATGCAAAAAGCAAAATG	1490
Db	146383	TGTACAGAGACAGGGATCGAGACCATCCCATGGAAAAGAAATGCAAAAAGCAAAATG	146324
QY	1491	GCTGTCTGGGAGGCTTACAAATAGCTGTGAAGAAGAGAGTGAAGAAGCAAGGAAA	1550
Db	146323	GCTGTCTGGGAGGCTTACAAATAGCTGTGAAGAAGAGAGTGAAGAAGCAAGGAAA	146264
QY	1551	AAAGGAAAGATAAAGCATCTGAATGACAGAGTCCAAA	1588
Db	146263	AAAGGAAAGATAAAGCATCTGAATGACAGAGTCCAAA	146204
QY	1589	-----	1588
Db	146203	AGCCTTCTCAGTGATCAATGCAAAAGAAATAGAGGAAAACACACAGAAATGGAAAGACTAG	146144
QY	1589	-----	1588
Db	146143	AGATCTCTCAAGAAATTAGAGATACCAAGGGGACATTTTCATGCAAGAGATGGCTCGAT	146084
QY	1589	-----	1588
Db	146083	AAAGGACAGAAATGATATGGACCTTAACAGAGCAGAGATATTAAAGAGAGATGGCAAGA	146024
QY	1589	-----	1588
Db	146023	ATACACAGAGAACTATACAAAAAGATCTTTCATGCCAGATATATCAGATGATGTGAT	145964
QY	1589	-----	1588
Db	145963	CACTGACCTAGAGCCAGACATCCTCGAATGGAATCAAGTGGGCTTTAGAAAGCATCAC	145904
QY	1589	-----	1588
Db	145903	TATGAACAAAGCCAGTGGAGGTGATGGAATTCAGTTGAGCTCTTCAATATCTGAAAGA	145844
QY	1589	-----	1588
Db	145843	TGATCTGTGAAAGTGTTCGACTCAATATGCCAGCAAAATTTGGAAAAACACAGCAGTGGCC	145784
QY	1589	-----	1588
Db	145783	ACAGGACTGGAAGAGTTCAGTTTTCATTCCTCAATCCCAAGAAAGCAANTGCCAAAGATG	145724
QY	1589	-----	1588
Db	145723	CTCAAACTACCGCACAAATTCGACTCATCTCGCATGCTAGTAAAGTAATGCTCAAAATTC	145664
QY	1589	-----	1617
Db	145663	CGAGCCAGGCTTCAGCAATACGTGAACCGTGAATTTCTGATGTTCAAGCTGGTTTTAG	145604
QY	1618	AAAGTTCAGAGAAACACAGAGCCAAATTCGCAACATCTCTGTATCATATGAAAAAGCAAG	1677
Db	145603	AAAGTTCAGAGAAACACAGAGCCAAATTCGCAACATCTCTGTATCATATGAAAAAGCAAG	145544
QY	1678	AGAGTTCAGAGAAACACATCTTCTGCTTTTATGACTATCAAAAGCTTTGACTGTGG	1737
Db	145543	AGAGTTCAGAGAAACACATCTTCTGCTTTTATTAACATATACCAAGCTTTGACTGTGG	145484
QY	1738	GGGTTCACATTAACCTGTGAAAAATTTCTGAAAGGGATGGAATACCAAGACCTGACCTG	1797
Db	145483	GGGTTCACATTAACCTGTGAAAAATTTCTGAAAGGGATGGAATACCAAGACCTGACCTG	145424
QY	1798	ACTCTTGAAGAAATTTGTATGAGGTTCAGAGTCAAGTGAAGTGAAGTGAAGTGAAGTGAAG	1857
Db	145423	ACTCTTGAAGAAATTTGTATGAGGTTCAGAGTCAAGTGAAGTGAAGTGAAGTGAAGTGAAG	145364
QY	1858	AGACTGGTTTCCAAAGTAGGAAAGGAGTATGTCAAGGCTGTATATTGTCAACCCGGCTTGT	1917
Db	145363	AGACTGGTTTCCAAAGTAGGAAAGGAGTATGTCAAGGCTGTATATTGTCAACCCGGCTTGT	145304
QY	1918	TAACTTCTATGAGAG-A-CATCATGAGAAACGCTGGCTGGAAGAGCAAGAGCTGGAAT	1976
Db	145303	TAACTTCTATGAGAGTACATCATGAGAAACGCTGGCTGGAAGAGCAAGAGCTGGAAT	145244
QY	1977	CAAGATTCCGGGAGAAATAGCAATTAACCTCAGATATGAGATATACCACTTATGGC	2036
Db	145243	CAAGATTCCGGGAGAAATAGCAATTAACCTCAGATATGAGATATACCACTTATGGC	145184
QY	2037	AGAAATGAAGAGAACTTAAAGAGCTTCTGATGAAGTGAAGAGGAGAGTGAAGAGT	2096
Db	145183	AGAAATGAAGAGAACTTAAAGAGCTTCTGATGAAGTGAAGAGGAGAGTGAAGAGT	145124
QY	2097	TGGCTTAAAGCTCAACATTCAGAAACCAAGATCATGTCATCTGTCCTCCATCATCTCATG	2156
Db	145123	TGGCTTAAAGCTCAACATTCAGAAACCAAGATCATGTCATCTGTCCTCCATCATCTCATG	145064
QY	2157	GGAAATAGATGGGAAACAGTGGAAACAGTGTGAGACTTTATTTTGGGGGGCTCCAAA	2216
Db	145063	GGAAATAGATGGGAAACAGTGGAAACAGTGTGAGACTTTATTTTGGGGGGCTCCAAA	145015


```
QY 2217 ATCACTGAGATGGTACTGACCCATGAATTAAGAAGACACTTACTCTTGGAGAAAA 2276
Db 145014 CTCAGTGCAGATGGTACTGACCCATGAATTAAGAAGACACTTACTCTTGGAGAAAA 144955
QY 2277 GTTA--ACCAACCTAGATACATATGAAGAAGACAGACATTAACCTTGGCAACAAGGCC 2334
Db 144954 GTTATGACCAACCTAGATGATATTAAGAAGACAGACATTAACCTTGGCAACAAGGCC 144895
QY 2335 CATCTAGTCAAGCTATGGTGTTCAGTGGTCTATGATGATGATGATGATGATGATGATG 2394
Db 144894 CTCTAGTCAAGCTATGGTGTTCAGTGGTCTATGATGATGATGATGATGATGATGATG 144835
QY 2395 AAGAAGCTGAGCACTGAAGAATGATGCTTTTGAAGTGGTGGTGGAGAGACTCTTG 2454
Db 144834 AAGAAGCTGAGCACTGAAGAATGATGCTTTTGAAGTGGTGGTGGAGAGACTCTTG 144775
QY 2455 AGAGTCCCTTGGCTGCAAGAGATCAACAGTCCATCTGAAGAGATCAGCCCTGGG 2514
Db 144774 AGAGTCCCTTGGCTGCAAGAGATCAACAGTCCATCTGAAGAGATCAGCCCTGGG 144715
QY 2515 ATTCTTTGGAAGGAATGATGCTAAAGCTGAAGTCACTTTGGCCACCTGATCAGA 2574
Db 144714 ATTCTTTGGAAGGAATGATGCTAAAGCTGAAGTCACTTTGGCCACCTGATCAGA 144655
QY 2575 AGAGTCACTCACTGGAAGAGCCCTGATCTGGAGGATTCGGGCGAGGAGAGAGG 2634
Db 144654 AGAGTCACTCACTGGAAGAGCCCTGATCTGGAGGATTCGGGCGAGGAGAGAGG 144595
QY 2635 GGACGACAGAGGATGAGATGGCTGGATGGGATCAGTCACTGATGATGATGATGATGATG 2694
Db 144594 GGACGACAGAGGATGAGATGGCTGGATGGGATCAGTCACTGATGATGATGATGATG 144535
QY 2695 GAATCTCTGAGTGGTGTGATGAGCAGGAGGCTGCTCGGCGATTCATGGGGTCA 2754
Db 144534 GAATCTCTGAGTGGTGTGATGAGCAGGAGGCTGCTCGGCGATTCATGGGGTCA 144475
QY 2755 AAGAGTGGACACGACTGAGCACTGAGCACTGAACTGAACT 2794
Db 144474 AAGAGTGGACACGACTGAGCACTGAGCACTGAACTGAACT 144435

RESULT 20
LOCUS AC144498/2 294688 bp DNA linear HTG 02-OCT-2003
DEFINITION Bos taurus clone RP42-25714, WORKING DRAFT SEQUENCE, 9 ordered
pieces.
ACCESSION AC144498
VERSION AC144498.3 GI:37360989
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoides;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 294688)
AUTHORS Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Cariaga,K.,
Chu,G., Coleman,B., Coleman,H., Engle,J., Granite,S., Guan,X.,
Gupta,J., Haghghi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P.,
Hurle,B., Idol,J.R., Karlins,E., Kwong,P., Lalic,P., Lee-Jin,S.-Q.,
Legaspi,R., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Mastaglio,C.,
Maskeri,B., McDowell,J., Padurigan,C., Pearson,R., Portnoy,M.E.,
Prasad,A., Reddix-Dugue,N., Schardier,K., Schueler,M.G., Shah,K.,
Simon,C., Stantripop,S., Thomas,J.W., Thomas,P.J., Tsipouri,V.,
Vest,J.L., Wetherby,K.D., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
REFERENCE 2 (bases 1 to 294688)
AUTHORS Green,E.D.
DIRECT SUBMISSION
TITLE Direct Submission
JOURNAL Submitted (25-APR-2003) NIH Intramural Sequencing Center, 8717
Grovenmont Circle, Gaithersburg, MD 20877, USA
REFERENCE 3 (bases 1 to 294688)
```

AUTHORS
TITLE
JOURNAL
COMMENT

Green,E.D.
Direct Submission
Submitted (02-OCT-2003) NIH Intramural Sequencing Center, 8717
Grovenmont Circle, Gaithersburg, MD 20877, USA
On Oct 2, 2003 this sequence version replaced gi:31621329.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@hgr.nih.gov
----- Project Information
Center project name: eem
Center clone name: 257104

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 293115 bases at least Q40
Consensus quality: 293600 bases at least Q30
Consensus quality: 293797 bases at least Q20
Insert size: 187000; agarose-fp
Quality coverage: 17.56x in Q20 bases; agarose-fp
Quality coverage: 11.18x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 9 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and the accession number will be preserved.

1
1193 1192: contig of 1192 bp in length
1293 1292: gap of unknown length
52165 52164: contig of 50872 bp in length
52265 52264: gap of unknown length
71132 71131: contig of 18867 bp in length
71232 71231: gap of unknown length
77082 77081: contig of 5850 bp in length
81754 81753: gap of unknown length
81854 81853: gap of unknown length
106710 106709: contig of 24856 bp in length
130077 130076: contig of 23267 bp in length
130177 130176: gap of unknown length
282833 282832: contig of 152656 bp in length
282933 282932: gap of unknown length
282933 282931: contig of 11756 bp in length.

FEATURES
source

1..294688
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="RP42-25714"
/clone_lib="RP42"
1..1192
/note="assembly_fragment"

misc_feature

```

clone_end:SP6
vector_side:left
misc_feature 1293..52164
  /note="assembly_fragment"
misc_feature 52265..71131
  /note="assembly_fragment"
misc_feature 71232..77081
  /note="assembly_fragment"
misc_feature 77182..81753
  /note="assembly_fragment"
misc_feature 81854..106709
  /note="assembly_fragment"
misc_feature 106810..1130076
  /note="assembly_fragment"
misc_feature 130177..282832
  /note="assembly_fragment"
misc_feature 239910..294688
  /note="clone overlaps with GenBank Accession Number
  AC144602 clone RP42-212N23 (center project name een)"
misc_feature 282933..294688
  /note="assembly_fragment
clone_end:right
vector_side:right

Query Match      57.8%; Score 2066.1; DB 1; Length 294688;
Best Local Similarity 76.5%; Pred. No. 0;
Matches 2544; Conservative 0; Mismatches 179; Indels 601; Gaps 7;

QY 126 TGGTAGGTGTTGTGAGAGGGGATCAGAGGGGAGACNACTGAAACCATACACGGAGAAA 185
DB 85445 TCGCAGGTGTTGCGAGGAGGGATCAGACGGGAAAACACACTGAAACCATACTACAGAAA 85386
QY 186 CTAGTCAATCTAATCACACTAGGACCAAGCTTGTCTAATCAATGAACTAAGCCATG 245
DB 85385 CTAGTCAATCTAATCACACTAGGACCAAGCTTGTCTAATCAATGAACTAAGCCATG 85326
QY 246 CCCGTGGGGCAACCCAGATGGGAGGTCTATGTTGGAGAGATCTGACAGAAATGTGTCCA 305
DB 85325 CCAGTGGGGCAACCCAGATGGGAGGTCTATGTTGGAGAGATCTGACAGAAATGTGTCCA 85266
QY 306 CTGGAGAGGGAAT-GCAACCACTTCAGTATCTTCCTCTGAAACCCCATGAACAGTA 364
DB 85265 CTGGAGAGGGAATGGCAACCACTTCAGTATCTTCCTCTGAAACCCCATGAACAGTA 85206
QY 365 TGAAGAGGCAAAATCATAGGATCTGAAAGAGGAACCTCCAGGTCTAGTAGTGGCCCCAT 424
DB 85205 TGAAGAGGCTAATGATAGGATCTGAAAGAGAACTCCCAAGTCTAGTAGTGGCCCCAT 85146
QY 425 ATGCTACTGGAGATCAGTGGAGAAATTAATCCAGAAAGATGAAGATGGAGCCAAAGC 484
DB 85145 ATGCTACTGGAGATCAGTGGAGAAATTAATCCAGAAAGATGAAGATGGAGCCAAAGC 85086
QY 485 AAAAAGAAATACCCAGCTGTGATGTGATCTGTGTATATAAGCAAGTCCGATGCTGTAAG 544
DB 85085 AAAAAGAAATACCCAGCTGTGATGTGATCTGTGTATATAAGCAAGTCTGATGTGTAAG 85026
QY 545 AGCAATATTCATAGGAACCTTGAATGTCTAGGTCCATGAATCAAGGCAAAATGGAAGTGG 604
DB 85025 AGCAATATTCATAGGAACCTTGAATGTCTAGGTCCATGAATCAAGGCAAAATGGAAGTGG 84966
QY 605 TCAACCAAGAGATGGCAAGAGTGAATGTCAACATTTCTAGGAATCAGCGAATTAATGGA 664
DB 84965 TTAAGCAAGAGATGGCAAGAGTGAATGTCAACATTTCTAGGAATCAGCGAATTAATGGA 84906
QY 665 CTGAAATGGGTGAATTTAATCTAGATGACCAATATATCTACTCTGCGGCGAGGAATCCC 724
DB 84905 CTGAAATGGGTGAATTTAATCTAGATGACCAATATATCTACTCTGCGGCGAGGAATCCC 84846
QY 725 TCAGAGAAATGGAGTAGGATCATGTCTACAAAGAGTCCGAAATGCAAGTACTTGGAT 784
DB 84845 TCAGAGAAATGGAGTAGGATCATGTCTACAAAGAGTCCGAAATGCAAGTACTTGGAT 84786
QY 785 GCAGTCTCAAAAACGACAGAATGATCTCTGTTGTTGTTTCCAAAGGCAAACTTCAATATCA 844

```

```

DB 84785 GCAATCTCAAAATGACAGATGATCTCTGTTCAATTTCCAAAGCAAACTTCAATATCA 84726
QY 845 CAGTAATCCAAGTCTATGCCCCAACCCAGTAATGTGTAAGAAAGCTGAAAGTGAACGGTCTT 904
DB 84725 CAGTAATCCAAGTCTATGCCCCAACCCAGTAATGTGTAAGAAAGCTGAAAGTGAACGGTCTT 84666
QY 905 ATGAGACCTACAGACCTTTTAGAATTAACCCCAAAAGAGATGCTCTTCTCATTATAG 964
DB 84665 ATGAGACCTACAGACCTTTTAGAATTAACCCCAAAAGAGATGCTCTTCTCATTATAG 84606
QY 965 GGGACTTGAATGCAAAAGTAGGAAGCAAAACACCTGGAGTAAACAGGCAAAATTTGGCC 1024
DB 84605 GGGACTTGAATGCAAAAGTAGGAAGTAAAGCAAAACACCTGGAGTAAACAGGCAAAATTTGGCC 84546
QY 1025 TTGGAATACGGAATGAAGAGGGGCAAAAGACTTAATAGAGTTTGGCAAGAAATGCACTGG 1084
DB 84545 TTGGAATACGGAATGAAGAGGGTAAAGACTTAATAGAGTTTGGCAAGAAATGCACTGG 84486
QY 1085 TCATAGCAAAACCCCTCTTCCAAACCAACAGAGAGAGACTCTACACATGGACATCACCAG 1144
DB 84485 TCATAGCAAAACCCCTCTTCCAAACCAACAGAGAGAGACTCTACACATGGACATCACCAG 84426
QY 1145 ATGGTCAACACCGAAATCAGATTGATTATATTTTTCAGGCCAAAGATGGAGAAAGCTCTA 1204
DB 84425 ATGGTCAACATCGAAATCAGATTGATTATATTTTTCAGGCCAAAGATGGAGAAAGCTCTA 84366
QY 1205 TACAGTACGCAAAACCAAGACAGGCTTACTGTGGCTCAGATCATCACTTCTTATTG 1264
DB 84365 TACAGTACGCAAAACCAAGACAGGCTTACTGTGGCTCAGATCATCACTTCTTATTG 84306
QY 1265 CCAAAATTCAGACTTAAATTTGAAGAAAGTAGGAAACCACTAGATCACTCAGGTAAAGACC 1324
DB 84305 CCAAAATTCAGACTTAAATTTGAAGAAAGTAGGAAACCACTAGATCACTCAGGTAAAGACC 84246
QY 1325 TAAATCCAAATCCCTTATGATTATACAGTGGAGTGAGAAATAGATTTAAGGGCCCTAGATC 1384
DB 84245 TAAATCCAAATCCCTTATGATTATACAGTGGAAATGAGAAATAGATTTAAGGGCCCTAGATC 84186
QY 1385 TCATAGACAGAGTACTCTAATGAATATATGGACAGAGGTTTCATGACATTTACAGAGACAG 1444
DB 84185 TCATAGATAGAGTCTGATGAACTATGGAATGAGTTTCATGACATTTACAGAGACAG 84126
QY 1445 GGATCGAGACCATCCCATGGAAAAGAAATGCAAAAAGCAAAATGGCTGTCTGGGGAGG 1504
DB 84125 GGATCAAGACCATCCCATGGAAAAGAAAGCAAAAAGCAAAATGGCTGTCTGGGGAGG 84066
QY 1505 CCTTACAAATAGCTGTGAAAAGAGAGAGTGAAGCAAAAGCAAAAGAGAAAGAGATATA 1564
DB 84065 CCTTACAAATAGCTGTGAAAAGAGAGAGTGAAGCAAAAGCAAAAGAGAAAGATATA 84006
QY 1565 AGCATCTGAATGACAGATTCCAAAG----- 1589
DB 84005 AACATCTGAATGACAGATTCCAAAGTAGTAAGAGAGATGAAGAAAGCTTCTTCTAGTG 83946
QY 1590----- 1589
DB 83945 ATCAATGCAAGAAATAGAGGAAACCAAGATGGGAAGAGACTAGGAATCTCTTCAAGA 83886
QY 1590----- 1589
DB 83885 AAATCAGAGATACCAAGGAACATTTTCATGCAAGATGAGCTCGATTAAGGACAGAAATG 83826
QY 1590----- 1589
DB 83825 GTATGGATCTAACAGAGCAGAGATATTAAGAAAGAGATGGCAAGAAATACACAGAGAAC 83766
QY 1590----- 1589
DB 83765 TGTAACAAAAGATCTTCCAGACCCAGATATACAGATGGTGTGGTCACTGACCTAGAGC 83706
QY 1590----- 1589

```

Db 83705 TAGAAATCCTGGAAATGTGAAGTCAAGTGGGCTTTAGAAAGCATCACTACAAACAAAGCTA 83646
Qy 1590 ----- 1589
Db 83645 GTGAGGTGATGGAATCCAACTGAGCTATTTCAAATCCTGAAAGATGATGCTGTGAAG 83586
Qy 1590 ----- 1589
Db 83585 TGCTGCACTCAATATGCCAGCAAAATTTGGAAACTCAGAGTGGCCACAGGATTTGGAATA 83526
Qy 1590 ----- 1589
Db 83525 GGTGAGTTTTCATTCCTCAATCCGGAAGAAAGCAATGCCAAGATGCTCAAACTACCGCA 83466
Qy 1590 ----- 1589
Db 83465 CAATTGCACTCATCTCAGATGCTAGTAAAGTAATGCTCAAAATTCCTCAAGCCAGGCTTC 83406
Qy 1590 ----- 1589
Db 83405 AGCAATATGTGAACCGTCAACTTCTGATGTTTCAAGCTGGTTTGGAAAGGACAGGAA 83346
Qy 1632 CCAGAGACCAAAATGCCAACAATCTCTGTATCATFGGAAAGCAAGAGAGTTTCCAGAAAA 1691
Db 83345 CCAGAGATCAAAATGCCAACAATCTCTGGATCATGGAAGAAAGCAAGAGAGTTTCCAGAAAA 83286
Qy 1692 ACATCTATTTCTGCTTTTATTTGACTATGCAAGGCTTTTGTACTGTGGGGTCACAATAAAC 1751
Db 83285 ACATCTATTTCTGCTTTTATTTGACTATGCAAGGCTTTTGTACTGTGGATCACAATAAAC 83226
Qy 1752 TGTGAAAAATCTGAAAGGATGGAATACAGACCACTGACCTGACTCTTCAAAAAAT 1811
Db 83225 TGTGAAAAATCTGAAAGAGATGGAATACAGACCACTGATCTGCCCCCTTGAGAAAT 83166
Qy 1812 TGATGAGGTGAGAGACCAAGTTAGACTGACATGGAACCAAGATGGAACCAAGATGTTTCAAG 1871
Db 83165 TGATGAGGTGAGAGACCAAGTTAGACTGACATGGAACCAAGATGGAACCAAGATGTTTCAAG 83106
Qy 1872 TAGAAAAAGGAGTATGCAAGGCTGTATTTGTCAAGGCTTTTGAAGTCTTATGCTATGAG 1931
Db 83105 TAGAAAAAGGATTCGTCAGGCTGTAGATGTCACCTGTTTATTTAATTTAAGCAG 83046
Qy 1932 AG-ACATCATGAGAAACGCTGGGTGGAAG-AAGCAACAAGTGAATCAAGATTTGCGGG 1989
Db 83045 AGTACATCATGAGAAACGCTGGGTGGAAGAAACCAAGCTGCAATCAAGATTTGCTGG 82986
Qy 1990 AGAATAGCAATACCTCAGATATGCAATGATACCACTTATGGCAGAAAGTGAAG 2049
Db 82985 AGAATATCAATAACCTTAGATATGCAATGATACCACTTATGGCAGAAAGTGAAG 82926
Qy 2050 GAACTAAAAAGCCTCTTGTATGAAGGTGAAGAGAGAGTGAAGAAAGTGGCTTAAAGCTC 2109
Db 82925 GAACTAAAAAGCCTCTTGTATGAAGGTGACAGTGAAGAGTGAAGAAAGTGGCTTAAAGCTC 82866
Qy 2110 AACATTCAGAAAAAGAGATCATGGATCTGGTCCCATCTATCATGGAAATAGATGGG 2169
Db 82865 AACATTCAGAAAAAGAGATCATGGATCTGGTCCCATCTATCATGGAAATAGATGGG 82806
Qy 2170 GAAACAGTGGAAACAGTGTGAGCTTTATTTTGGGGGCTCCAAATCACTGCAGATG 2229
Db 82805 GAAACAGTGGAAACAGTGTGAGCTTTA--TTTGGGGGCTCCAAATCACTGCAGATG 82748
Qy 2230 GTGACTGAGCCATGAAATTAAGACACTTATCTCTTGGAAAGAAAGTTA--ACCAACC 2287
Db 82747 GTGACTGAGCCATGAAATTAAGAGATGCTCTCTCTTGGAAAGAAAGTTATGACCAACC 82688
Qy 2288 TAGATGACATATTGAAAGAGAGACATTAACCTTGGCCAAAGAGCCCATCTAGTCAAGG 2347
Db 82687 TAGACAGATATTCAAAAGAGAGACATTAACCTTGGCCAAAGAGATGCTCTAGTCAAGG 82628
Qy 2348 CTATGGTTTTTCCAGTGGTCAATGATGATGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2407
Db 82627 CTATGGTTTTTCCAGTGGTCAATGATGATGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 82568

Qy 2408 ACTGAAGAATTGATGCTTTTGAACCTGTGTGTGGAGAGACTCTTTGAGAGTCCCTTGA 2467
Db 82567 GCCAAAGAATTGATGCTTTTGAACCTGTGTGTGGAGAGACTCTTTGAGAGTCCCTTGA 82508
Qy 2468 CTGCAAGGAGATCCCAACAGTCCATTCGAAGAGATCAGCCCTGGATTTCTTTGGAG 2527
Db 82507 CTGCAAGGAGATCCCAACAGTCCATTCGAAGAGATCAGCCCTGGATTTCTTTGGAG 82448
Qy 2528 GAATGATGCTAAAGTGTAAACCTCCAGTACTTTGGCCACTGATCAGAGAGCTGACTCAC 2587
Db 82447 GAATGATGCTAAAGTGTAAACCTCCAGTACTTTGGCCACTGATCAGAGAGTGTACTCAT 82388
Qy 2588 TGAAGAAGACCTGTGATGCTGGAGGATTTGGGGGAGAGAGAGAGAGAGAGAGAGAGAG 2647
Db 82387 TGAAGAAGACCTGTGATGCTGGAGGATTTGGGGGAGAGAGAGAGAGAGAGAGAGAG 82328
Qy 2648 TGAGATGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2707
Db 82327 TGAGATGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 82268
Qy 2708 TGGTATGATGAGAGGAGGCTCTCTCTGGGCGGCAATTCATGGGGTCAAAAGATTTGGACAC 2767
Db 82267 TGGTATGATGAGAGGAGGCTCTCTCTGGGCGGCAATTCATGGGGTCAAAAGATTTGGACAC 82208
Qy 2768 GACTGAGCAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACT 2827
Db 82207 AACTGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 82149
Qy 2828 AAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2851
Db 82148 GGTAAAGATGCTCTTCAATGCA 82125

RESULT 21

AC135216

LOCUS

AC135216 153264 bp DNA linear HTG 27-NOV-2002
Bos taurus clone RP42-427020, WORKING DRAFT SEQUENCE, 3 ordered pieces.

AC135216

AC135216.2 GI:25699761

HTG; HTGS_PHASE2; HTGS_DRAFT.

KEYWORDS

SOURCE

Bos taurus (cow)

ORGANISM

Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovidae; Bovinae; Bos.

1 (bases 1 to 153264)

Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,

Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,

Carliaga,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J.,

Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E.,

Laric,P., Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,

Margulies,E.H., Mastello,C., Maskeri,B., McDowell,J.,

Paguirigan,C., Pearson,R., Portnoy,M.E., Prasad,A.,

Reddy-Dugue,N., Schandler,K., Schueler,M.G., Sison,C.,

Stadler-Pop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.L.,

Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.

NISC Comparative Sequencing Initiative

Unpublished

2 (bases 1 to 153264)

Green,E.D.

Direct Submission

Submitted (09-OCT-2002) NIH Intramural Sequencing Center, 8717

Grovemont Circle, Gaithersburg, MD 20877, USA

3 (bases 1 to 153264)

Green,E.D.

Direct Submission

Submitted (27-NOV-2002) NIH Intramural Sequencing Center, 8717

Grovemont Circle, Gaithersburg, MD 20877, USA

On Nov 27, 2002 this sequence version replaced gi:23622199.

----- Genome Center

Center: NIH Intramural Sequencing Center

Qy	1199	GCTCTATACAGTCAGCAAAAAACAAGCAGGAGCTTACTGTGGCTCAGATCATGAATCC	1258
Db	148824	GCTCTATACAGTCAGCAAAAAACAAGCAGGAGCTGACTGTGGCTCAGATCATGAATCC	148883
Qy	1259	TTATTGCCAAATTCAGACTTAAATTTGAAGAAGTAGGAAAAACCACATAGATCACTCAGGT	1318
Db	148684	TTATTGCCAAATTCAGACTTAAATTTGACGAAGTAGGGGAAAACCACTAGACCAATTCAGT	148943
Qy	1319	AAGACCTAAATCCAAATCCCTTAATGATTATACAGTGGAAATGAGAAATAGATTTTAAGGGCC	1378
Db	148944	ATGACCTGAATCAAAATCCCTTAATGATTATATGTTGAAAGTCAGAAATAGATTTTAATGGAC	149003
Qy	1379	TAGATCTGATAGACAGCTACTATGAATGAATATGGACAGAGTTTCATGACATTTACAGG	1438
Db	149004	TAGATCTGATAGACAGCTGCTGTAAGCTATGGAATGAGTTTCGTGACATTTGTACAGG	149063
Qy	1439	AGACAGGGATCGAGACCATCCCAATGGAAAAAGAAATGCAAAAAAGCAAAATGGCTGTCTG	1498
Db	149064	GTACAGGGATCAAGACCATCCCAATGGAAAAAGAAATGCAAAAAAGCAAAATGGCTGTCTG	149123
Qy	1499	GGGAGCCCTTACAAATAGCTGTGAAAAAGAGAAAGTGAAGCAAGCAAGAAAAAGGAAA	1558
Db	149124	GGGAGGCCCTTATAAACAGCTGTGAAAAAGAGAAAGCGAAAAAGCAAGAAAAAGGAAA	149183
Qy	1559	GATAAAAGCATCTGAATGCAGAGTTCCAAA	1588
Db	149184	GATATAAGCATCTGAATGCAGAGTTCCAAAAGATAGCAAGAAGATAGAAGCCTTCTCTC	149243
Qy	1589	-----	1588
Db	149244	AGTGATCAATGCAAAAGAAATAGAGAAAAACAACAGAAATGGAAAGACTAGAGATCTCTTC	149303
Qy	1589	-----	1588
Db	149304	AAGAAATTTAGAGATACCCAGGGAACATTTTCATGCAAGAAGGGCTAGATAAAGGACAA	149363
Qy	1589	-----	1588
Db	149364	AATGGTATGGACCTAACAGAAGTAGAGATATTAAGAAGAGGTGGCAAGAATACACAGAA	149423
Qy	1589	-----	1588
Db	149424	GAATGTAACAAAAGATCTTCATGATCTAGATATCACAAATGGTGTGATCATGTACCTTA	149483
Qy	1589	-----	1588
Db	149484	GAGCCAGACATCCTGGAATGTGAAGTCAAGTGGGCTTTAGAAAGCATCACTACGAACAA	149543
Qy	1589	-----	1588
Db	149544	GCTAGTGGAGGTGATGGAATTCAGTTGAGCTATTTCCAAATCCTTAAAGATGATGTGTG	149603
Qy	1589	-----	1588
Db	149604	AAAGTGTGCACTCAATATGCCAGCAATTTTGGAAAACTCAGCAGTGGCCACAGGACTGG	149663
Qy	1589	-----	1588
Db	149664	AAAAAGTCAATTTCAATTCCAATTCCTCAAAAAAGGCAATGCCAAAGATGCTCAAACTAC	149723
Qy	1589	-----	1588
Db	149724	ATCAAAATTTCAATCTCACAGGCTAGTAAGTAATGCTCAAAATTTCTCAAGCCAGGC	149783
Qy	1589	-----	1628
Db	149784	TTACAGCAATATGTGAACCATGAATTCAGATGTTTCAAGCTGGTTTTAGAAAAAGGACAG	149843
Qy	1629	GAACAGAGCAAAATTCACACATCTCTGATCATGGAAGCAAGCAGAGTTCCAGA	1688
Db	149844	GAACAGAGATCAAAATTCGCAACATCTCTGTGTAATGGAAGCAAGAGTTCCAGA	149903

Qy	1689	AAAA	CATCTATTTC	TGCTTTTAT	TGAC	TAATG	CAAA	AGCCTTT	GACTGTGGGGT	CAACA	TA	1748
Db	149904	AAAA	CATCTATTTC	TGCTTTTAT	TGAC	TAATG	CAAA	AGCCTTT	GACTGTGGGGT	CAACA	TA	149963
Qy	1749	AACT	GTGGAAAA	TTCT	GAAAGGGAT	GGA	ATAC	CAGAC	CACTGACCT	GACTCTTT	GAAAA	1808
Db	149964	AACT	GTGGAAAA	TTCT	GAAAGGGAT	GGA	ATAC	CAGAC	CACTGACCT	GACTCTTT	GAGAA	150023
Qy	1809	ATTT	TGATG	CAGGTC	AGGA	ACCA	CAGTT	AGAA	CACTGGA	CAACAG	ACTGGTTCC	1868
Db	150024	ACCT	ATATG	CAGGTC	AGGA	ACCA	CAGTT	AGAA	CACTGGA	CAACAG	ACTGGTTCC	150083
Qy	1869	AAGT	AGGAAAA	AGGAGT	ATGTC	CAAG	CGCTAT	TATGTC	CACCGCGT	TGTTAA	CTTCTATG	1928
Db	150084	AAAT	AGGAAAA	AGGAGT	ATGTC	CAAG	CGCTAT	TATGTC	CACCGCGT	TATTA	CTTCTATG	150143
Qy	1929	CAGAG	-ACAT	CATG	AGAA	AC	CTGG	CTGGA	AGAC	CAACAG	CTGGAATCA	1987
Db	150144	CAGAGT	ACAT	CATG	AGAA	AC	CTGG	CTGGA	AGAC	CAACAG	CTGGAATCA	150203
Qy	1988	GGAGA	ATAG	CAAT	TAAC	CTC	AGATAT	GCAGAT	GTATAC	CACCCCTT	ATGCGAGAAAGTGA	2047
Db	150204	GGAGA	ATAT	CCAT	TAAC	CTC	AGATAT	GCAGAT	GTATAC	CACCCCTT	ATGCGAGAAAGTGA	150263
Qy	2048	AGGA	CTAAAA	AGCCTCT	TGAT	GAAG	GTGAA	AGGAG	GAGGTG	AAAAAGT	TGCGCTTAAAGC	2107
Db	150264	AGGA	CTAAAA	AGCCTCT	TGAT	GAAG	GTGAA	AGGAG	GAGGTG	AAAAAGT	TGCGCTTAAAGC	150323
Qy	2108	TCA	CAATTC	CAGAAA	AC	GNAG	ATCAT	GGCAT	CTGGTCC	CATCAC	TTTCATGGG	2167
Db	150324	TCA	CAATTC	CAGAAA	AC	GNAG	ATCAT	GGCAT	CTGGTCC	CATCAC	TTTCATGGG	150383
Qy	2168	GGG	AAACAGT	GAAA	ACAGT	GT	CAG	ACTTTA	TTTTTT	TGGGGGG	CTCAAAAT	2227
Db	150384	GGG	AAACAGT	GAAA	ACAGT	GT	CAG	ACTTTA	TTTTTT	TGGGGGG	CTCAAAAT	150441
Qy	2228	TGGT	GACTG	CAG	CCAT	TGAA	ATTTAA	AAAG	CAC	TTACTCT	CTTGGAGAA	2285
Db	150442	TGGT	GATTG	CAG	CCAT	TGAA	ATTTAA	AAAG	CAC	TTACTCT	CTTGGAGAA	150501
Qy	2286	CCT	AGATAG	CAT	ATTG	AAAA	AC	CAGAC	ATTAC	CTTG	CCAA	2345
Db	150502	CCT	AGATAG	CAT	ATTG	AAAA	AC	CAGAC	ATTAC	CTTG	CCAA	150561
Qy	2346	GGCT	ATGGTTTT	TTCC	AGTGGT	CAT	GTAT	GATG	TGAG	AGTTGG	ACTGTGA	2405
Db	150562	GGCT	ATGGTTTT	TTCC	AGTGGT	CAT	GTAT	GATG	TGAG	AGTTGG	ACTGTGA	150621
Qy	2406	GCA	CTGA	AGAA	TTGAT	CTG	TTT	GA	CTGTG	TGAG	AGACTCT	2465
Db	150622	GTG	CCAA	GA	AAATGAT	CTG	TTT	GA	CTGTG	TGAG	AGACTCT	150681
Qy	2466	GACT	GC	AA	GAGAG	ATCC	AC	AGTCC	ATTCT	GA	AGAGAT	2525
Db	150682	GACT	GC	AA	GAGAG	ATCC	AC	AGTCC	ATTCT	GA	AGAGAT	150741
Qy	2526	AGGA	ATGAT	GTAA	AGCT	GAA	ATCT	CCAGT	ACTTT	TGGCC	ACCTGAT	2585
Db	150742	AGGA	ATGAT	GTAA	AGCT	GAA	ATCT	CCAGT	ACTTT	TGGCC	ACCTGAT	150801
Qy	2586	ACT	GG	AAAA	AGAC	CC	CTGAT	CTG	GGAGGA	-TTG	GGGG	2644
Db	150802	ATT	GG	AAAA	AGAC	CC	CTGAT	CTG	GGAGGA	-TTG	GGGG	150861
Qy	2645	GGAT	GAGAT	GCT	GGAT	TGG	CA	TG	ACT	CGAT	TGGA	2704
Db	150862	GGAT	GAGAT	GCT	GGAT	TGG	CA	TG	ACT	CGAT	TGGA	150921
Qy	2705	AGTT	GGT	GAT	GG	AC	AGG	AGC	CTCT	CTCG	GGGAT	2764
Db	150922	AGTT	GGT	GAT	GG	AC	AGG	AGC	CTCT	CTCG	GGGAT	150981
Qy	2765	CAC	GACT	GA	CA	ACT	GA	CTGA	CTGA	2786		

```

Db 150982 CATGACTGAGCGACTGAAGTGA 151003
|||||
AC147195 223761 bp DNA linear HTG 17-DEC-2003
Bos taurus clone RP42-466M15, WORKING DRAFT SEQUENCE, 4 ordered
pieces.
AC147195
AC147195.2 Gi:39979451
HTG; HTGS PHASE2; HTGS_DRAFT.
SOURCE
Bos taurus (cow)
ORGANISM
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 223761)
Antonellis,A., Ayele,K., Benjamin,B., Blakesley,R.W.,
Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,B.,
Coleman,H., Daki,N., Engle,J., Granite,S., Guan,X., Gupta,J.,
Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P., Hurle,B.,
Iqbal,J.R., Jones,C., Karlins,E., Kim,H., Kwong,P., Lalic,P.,
Larson,S., Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,
Marquis,E.H., Masiello,C., Maskeri,B., McDowell,J.,
Mullikin,J.C., Naguirigan,C., Portnoy,M.E., Prasad,A., Puri,O.,
Reddik-Dugue,N., Schandler,K., Schueler,M.G., Shah,K., Sison,C.,
Stantrop,S., Thomas,J.W., Thomas,P.J., Tsipouri,V., Vogt,J.L.,
Wetherby,K.D., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
REFERENCE
2 (bases 1 to 223761)
Green,E.D.
Direct Submission
Submitted (08-NOV-2003) NIH Intramural Sequencing Center, 8717
Government Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 223761)
Green,E.D.
Direct Submission
Submitted (17-DEC-2003) NIH Intramural Sequencing Center, 8717
Government Circle, Gaithersburg, MD 20877, USA
On Dec 17, 2003 this sequence version replaced gi:38194246.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@hghri.nih.gov
----- Project Information
Center project name: ees
Center clone name: 466M15

The sequence data in this record represents an 'enhanced'
version of a Phase 2 submission. Specifically, the indicated
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8X average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associated
with a Phrap-derived quality score.
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 223320 bases at least Q40
Consensus quality: 223438 bases at least Q30
Consensus quality: 223450 bases at least Q20
Insert size: 178000; agarose-fp
Insert size: 223461; sum-of-contigs
Quality coverage: 27.55x in Q20 bases; agarose-fp

```

Quality coverage: 21.94x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.

1 145383: contig of 145383 bp in length

145384 171206: contig of 25723 bp in length

145484 171306: gap of unknown length

171207 209669: contig of 38363 bp in length

171307 209670: gap of unknown length

209670 223761: contig of 13992 bp in length.

209770 223761: contig of 13992 bp in length.

Location/Qualifiers

1. 223761

/organism="Bos taurus"

/mol_type="genomic DNA"

/db_xref="taxon:9913"

/clone="RP42-466M15"

/clone_lib="RP42"

/note="BAC resource: http://bacpac.chori.org/"

1. 145383

/note="assembly_fragment"

clone_end:SP6

vector_side:left

145484. 171206

/note="assembly_fragment"

171307. 209669

/note="assembly_fragment"

209770. 223761

/note="assembly_fragment"

clone_end:17

vector_side:right

misc_feature

1. 145383

/note="assembly_fragment"

clone_end:SP6

vector_side:left

145484. 171206

/note="assembly_fragment"

171307. 209669

/note="assembly_fragment"

209770. 223761

/note="assembly_fragment"

clone_end:17

vector_side:right

Query Match 57.8%; Score 2065; DB 1; Length 223761;

Best Local Similarity 76.7%; Pred. No. 0;

Matches 2561; Conservative 0; Mismatches 165; Indels 614; Gaps 9;

QY 81 GTAAAGAGATACCCACGCCCGAAGGTAGAGAAACCCAGTAAGATGTTGTGTA 140

DB 198857 GTGAGGAATATACCGTCGTCCAAAGGTAGAGAAACCCAGTAAGATGTTGCA 198916

QY 141 GAGGGCATCAGAGGGGAGACATCTGAAACCATACACGAGAAACCTAGTCAATCTAATC 200

DB 198917 GAGGGCATCAGAGGGGAGACACACCTGAAACCATCTACAGAAACTAGTCAATCTAATC 198976

QY 201 ACATAGGACCCACAGCCTTGTCTAACTCAATGAATGAAGCATGCCGTGGGCAACCC 260

DB 198977 ACATAGGACCCACAGCCTTGTCTAACTCAATGAATGAAGCATGCCGTGGGCAACCC 199036

QY 261 AGATGGGCGAGTCATGTTGGAGAGATCTGACAGATGTGTCCACTGGAGAGGAAT- 319

DB 199037 AGACGGGCGAGTCATGTTGGAGAGAACTGACAGATGTGTCTACTGGAGAGGAATG 199096

QY 320 GCAAAACCACTTCAGTATTCTTTCCTTGGAGAACCCCATGAACAGTATGAAAGGCAAAATG 379

DB 199097 GCAAAACCACTTCAGTATTCTTTCCTTGGAGAACCCCATGAACAGTATGAAAGGCAAAATG 199147

QY 380 ATAGATATCTGAAGAGGAACCTCCCGCTAGTGTGCCCCATCTACTCTGAGATC 439

DB 199148 ATAGATATCTGAAGAGGAACCTCCCGCTAGTGTGCCCCATCTACTCTGAGATC 199207

QY 440 ACTGGAGAAATTAATCCGAAAGAAATGAAGAGATGAGCCCAAGCAAAAGAAATCCAG 499

DB 199208 AGTGAGAGAAATTAATCCGAAAGAAATGAAG-GACGAGCCCAAGCAAAAGCAATACCCAA 199266

QY 500 CTGTGATGTGACTGGTGATATAACCAAGTCCCGTCTGTAAAGAGCAATATTCATAG 559

D	b	199267	CTGTGGATGTGACTGGTAAATAGAAAGCAAGTCTGATGCTCTAAAGAGCAATATTTGCATAG	199326	Q	y	1589	-----	1589
Q	y	560	GAACCTGGAATGTGACAGTCCATGAATCAAGGCCAAATTTGGAAGTGGTCAAAACAAGAGATGG	619	D	b	200407	ACAAATAGAGGAAAAACAACAGAAATGAGAAAGACTAGAGATCTCTTCAAGAAAAATCAAGAT	200466
D	b	199327	GAACCTGGAATGTGACAGTCCATGAATCAAGGCCAAATTTGGAAGTGGTCAAAACAAGAGATGG	199386	Q	y	1589	-----	1588
Q	y	620	CAAGAGTGAATGTCAACATTTCTAGGAATCAGCAATCAAAATGGAATGGATGGTGGTGAAT	679	D	b	200467	ACCAAAGGAACATTTCTATGCAAAAGATGGCTTGATAAAGGACAGAAATGGTATGACCTA	200526
D	b	199387	GAAGAGTGAATGTGCAATTTCTAGGAATCAGCAATCAAAATGGAATGGATGGTGGTGAAT	199446	Q	y	1589	-----	1588
Q	y	680	TTAACTCAGATGACCATATATATCTACTACTCGCGGACAGGAATCCCTCAGAGAAATGGAG	739	D	b	200527	ACAGAAGCAGAAGATATTAAAGAAGAGATGGCAATAATACACAGAGAAACTGTACAAAAA	200586
D	b	199447	TTAACTCAGATGACCATATATATCTACTACTCGCGGACAGGAATCCCTCAGAGAAATGGAG	199506	Q	y	1589	-----	1588
Q	y	740	TAGCCATCTAGTCCAAACAAGAGTCCGAAATGCAAGTACTTTGATGCAAGTCTCAAAAACG	799	D	b	200587	GATCTTCAACCCAGATAATCACGATGGTGTGATCACTGACCTTAGAGCCAGACATCCTG	200646
D	b	199507	TGGCCATCATGGTCAACAAAGAGTCTGAAATGCAAGTACTTTGATGCAATCTCAAAAATG	199566	Q	y	1589	-----	1588
Q	y	800	ACAGAATGATCTCTGTTTGTTCAGGCCAAACCATTCATATCACAGTAATCCAAAGTCT	859	D	b	200647	GAATGTGAAGTCAAGTGGGCCCTTTAGAAAGCATCACTACGAACAAAGCTAGTGGAGGTGAT	200706
D	b	199567	ACAGAATGATCTCTGTTTGTTCAGGCCAAACCATTCATATCACAGTAATCCAAAGTCT	199626	Q	y	1589	-----	1588
Q	y	860	ATGCCCAACAGGPAATGCTGAAGAGCTGAAGTTGAACGGTCTCTATGAAGACTTCAAG	919	D	b	200707	GGATTTCCAGTTGAGCTATTCCAAATCCCTGAAAGATGATGCTGTGAAAAGTCTGCACCTCA	200766
D	b	199627	ATGCCCAACAGGPAATGCTGAAGAGCTGAAGTTGAACGGTCTCTATGAAGACTTCAAG	199686	Q	y	1589	-----	1588
Q	y	920	ACCTTTTAGAAGTCAACACCCAAAGAGATGCTCTCTCTCATTTATAGGGACTGGAATGCAA	979	D	b	200767	ATATGCCACACATTTTGGNAAACTCAGCAGTGGCCACAGGACTAGAAAAGGTCAAGT	200826
D	b	199687	ACCTTTTAGAAGTCAACACCCAAAGAGATGCTCTCTCTCATTTATAGGGACTGGAATGCAA	199746	Q	y	1589	-----	1588
Q	y	980	AAGTAGGAAGCAAGAACACCTGGAGTAAACAGCAAAATTTGGCTTTGGAATCGGAATG	1039	D	b	200827	ATTCCAATCCCAAGAAAGGCAACGCCAAAGAAATGCTCAAACTACTGCACAAATTCGACTC	200886
D	b	199747	AAGTAGGAAGTCAAGAACACCTGGAGTAAACAGCAAAATTTGGCTCTGTAATCGGAATG	199806	Q	y	1589	-----	1588
Q	y	1040	AAGCAGGCCAAGACTAATAGAGTTTCCAGNAATGACCTGGTCAATAGCAACACCC	1099	D	b	200887	ATCTCACACACTAGTAAAGTAATGCTCAACATTTCCAAAGCCAGGCTTCAGCAATATCTG	200946
D	b	199807	AAGCAGGCCAATAGACTAATAGAGTTTCCAGNAATGACCTGGTCAATAGCAACACCC	199866	Q	y	1589	-----	1642
Q	y	1100	TCCTCCAAACACACAGAGAGACTCTACACATGGACATCACAGATGGTCAACACCGAA	1159	D	b	200947	AACCATGAACCTCTGTGATGTTTCAAGCTGTTTACAGCTGTTTACAGAGGACAGAGTGC	201006
D	b	199867	TCCTCCAAACACACAGAGAGACTCTATACATGGACATCACAGATGGTCAACACTGAA	199926	Q	y	1643	ATTGCCAATCCTCTGTATCATGGAAGAACAGAGAGTCCAGAAAAACATCTATTTC	1702
Q	y	1160	ATCAGATTGATATATTTCTTGGAGCAAGATGGAGAGCTCTATACAGTCAAGCAAAA	1219	D	b	201007	ATTGCCACCATCTGTGGATCATGSCAAAGCAAGAGAGTCCAGAAAAGCATCTATTTC	201066
D	b	199927	ATCAGATTGATATATTTCTTGGAGCAAGATGGAGAGCTCTATACAGTCAAGCAAAA	199986	Q	y	1703	TGCTTTATTGACTATGCAAAAAGCCTTTGACTGTGGGGTCAAAATAAACTGTGGAATAAT	1762
Q	y	1220	CAAGCAGGAGCTTACTGTGGCTCAGATCATGAATCTCTTATGTCCTTATGCAATTCAGACTTA	1279	D	b	201067	TGCTTTATTGACTATGCCAAAGCCTTTGACGTGTGGATCAAAATAGCTGTGGHAACT	201126
D	b	199987	CAAGCAGGAGCTTACTGTGGCTCAGACCATGAATCTCTTATGTCCTTATGTCCTTATG	200046	Q	y	1763	CTGAAAGGATGGGAATACACAGCAACCTGACCTGACCTCTTGAAAAATTTGTATGCAGGT	1822
Q	y	1280	AATTGAAGAAAGTAGGGAACCACTAGATCACTCAGGTAAGACCTAAATCCAAATCCCTT	1339	D	b	201127	CTGAAAGAGATGGGAATACACAGCAACCTGATCTGCTCTTGAGAAATTTGTATGCAGGT	201186
D	b	200047	AATTGAAGAAAGTAGGGAACCACTAGATCACTCAGGTAAGACCTAAATCCAAATCCCTT	200106	Q	y	1823	CAGGAAGCAACAGTTAGAACCTGGACATGGAACCAACAGACTGGTTCCTCAAGTAGGAAAAAG	1882
Q	y	1340	ATGATTATACAGTGGAGTGAATAGATTAAAGGCTTAGATCTGATA-----GACAG	1394	D	b	201187	CAGGAAGCAACAGTTAGAACCTGGACATGGAACCAACAGACTGGTTCCTCAAGTAGGAAAAAG	201246
D	b	200107	ATGATTATACAGTGGAGTGAATAGATTAAAGGCTTAGATCTGATA-----GACAG	200166	Q	y	1883	GTATGTCAAGGCTGTATATTGTCACCCGCTTTGTTAACTTCTATGTCAGAG-ACATCATG	1941
Q	y	1395	AGTACCTAATGAACTATGGACAGAGGTTCTATGACATTTGTAAGAGACAGGATTCAGAC	1454	D	b	201247	GTTCGTCAAGGCTGTATATTGTCACCCCTGCTTATTAACTTATATGTCAGAGTACATCATG	201306
D	b	200167	AGTACCTAATGAACTATGGACAGAGGTTCTATGACATTTGTAAGAGACAGGATTCAGAC	200226	Q	y	1942	AGAAACGCTGGCTGGAGAGAACCAACGCTGGAATCAAGATTGCCGGGAGGAATAGCAAT	2001
Q	y	1455	CATCCCATGAAAGAAATGCAAAAAGCAAAATGCTGTCTGGGGAGGCTTACAAAT	1514	D	b	201307	AGAAACGCTGGACTGGAGAGAACCAACGCTGGAATCAAGATTGCCGGGAGGAATATCATG	201366
D	b	200227	CATCCCATGAAAGAAATGCAAAAAGCAAAATGCTGTCTGAGAGGCTTACAAAT	200286	Q	y	2002	AACCTCAGATATGTCAGATGATACCAACCTTATGGCAGAAAGTGAAGAGGAACCTAAAAAGC	2061
Q	y	1515	AGCTGTGAAGAGAGAGAGTGAAGACCAAGGAAAGAAAGAGAGATATAAGAGATCTGAA	1574	D	b	201367	AACCTCAGATATGTCAGATGACACACCTTATGGCAGAAAGTGAAGAGGAACCTCAAGC	201426
D	b	200287	AGCTGTGAAGAGAGAGTGAAGACCAAGGAAAGAGAGATATAAGAGATATAAGATCTGAA	200346	Q	y	2062	CTCTTGATGAAGGTGAAGAGGAGAGTGAAGAGTGGCTTAAAGCTCAACATTCAGAAA	2121
Q	y	1575	TGCAGAGTTCAAA-----	1588	D	b	201427	CTCTTGATGAAGGTGAAGAGTGAAGAGTGGCTTAAAGCTCAACATTCAGAAA	201486
D	b	200347	TGCAGAGTTCAAAAGAAATAGCAAGAGAGATAGAAAGGCTTCTTTCAGCGATCAATGCAA	200406					


```
/clone="RP42-118011"
/clone_lib="RP42"
1. 10238
/clone="assembly_fragment"
vector_side:left
10339. 20461
/clone="assembly_fragment"
20562. 41485
/clone="assembly_fragment"
41586. 43195
/clone="assembly_fragment"
43296. 51664
/clone="assembly_fragment"
51765. 140032
/clone="assembly_fragment"
140133. 180432
/clone="assembly_fragment"
clone_end:SP6
vector_side:right

Query Match      57.7%; Score 2062.1; DB 1; Length 180432;
Best Local Similarity 74.7%; Pred. No. 0;
Matches 2568; Conservative 0; Mismatches 213; Indels 655; Gaps 8;

Qy 1 GTCAGGAGGCGCGCAGTGAAGAGGTACTACCTCGTCCAGGTGAAGAGCAGTAGC 60
Db 42506 GTCAGGAGGCGCGGTGAGGAGA-----TACCCCTCGTCCAGGTGAAGAGCAATGGC 42560
Qy 61 TGCCTTTGCTGGAGCAGCGCTAAAGAGATACCCACGCCACCAAGGTAAAGAGAAACCCAAAG 120
Db 42561 TGCCTTTGCTGGAGCAGCGGTGAAGAGATACCCACGCCACCAAGGTAAAGAGAAACCCAAAG 42620
Qy 121 TAAGATGGTAGTGTGTGAGAGGCGCATCAGAGGCGAGACATCTGTAAGAACCATACAGCA 180
Db 42621 TAAGATGGTAGTGTGCAAGAGGCGCATCAGAGGCGCAACACACTGAAACCATACTCACA 42680
Qy 181 GAAACTAGTCAATCTAATCACACTAGGACACAGCCTCTCTAATCTCAATGAATTAAG 240
Db 42681 GAAACTAGTCAATCTAATCACACTAGGACACAGCCTCTCTAATCTCAATGAATTAAG 42740
Qy 241 CCATGCCCGTGGGCAACCCAAAGATGGGCGAGTCACTGGTGAGAGATCTGACAGAAATGTG 300
Db 42741 CCATGCCCGTGGGCAACCCAAAGACGGCGGTCACTGGTGAGAGATTTGATAGAAATGTG 42800
Qy 301 GTCCACTGGAGAGGGGAAAT-GCAAAACCACTTCAGTATTTCTGCTTGAGAACCCCATGAA 359
Db 42801 GTCCACTGGAGAGGCAATGGCAAAACCACTTCAGTATTTCTGCTTGAGAACCCCATGAA 42860
Qy 360 CAGTATGAAAGGCAAAATGATAGGATACGAAAGAGAACTCCCGCAGGTCACTAGTGTGC 419
Db 42861 CAGTATGAAAGGCAAAATGATAGGATACGAAAGAGAACTCTCAGGTCACTAGTGTGC 42920
Qy 420 CCCATATGCTACTGGAGATCAGTGGAGAAATACTCCAGAAAGAAATGAAAGATGGAGCC 479
Db 42921 CCAATATGCTACTGGAGATCAGTGGAGAAATACTCCAGAAATAATGAAGGGATGGAGCC 42980
Qy 480 AAGCBAAGAAATACCCAGCTGGATGTGACTGTGATATAGCAAGTCCGATGCTG 539
Db 42981 AAGCBAAGAAATACCCAGCTGGATGTGACTGTGATATAGCAAGTCCGATGCTG 43040
Qy 540 TAAAGAGCAATATTGTCATAGGAACCTGGAATGTGAGGTCCATGAATCAAGGCAAAATTGGA 599
Db 43041 TAAAGAGCAATATTGTCATAGGAACCTGGAATGTGAGGTCCATGAATCAAGGCAAAATTGGA 43100
Qy 600 AGTGGTCAACAGAGATGGCAAGAGTGAATGTCAACATCTAGGAATCAGCGAACTAAA 659
Db 43101 AGTGGTCAACAGAGATGGCAAGAGTGAATGTGTCGACAGTCTAGGAATCAGCAAACTGAA 43160
Qy 660 ATGACCTGGATGGGTGAATTTAACTCAGATGACC----- 694
Db 43161 ATGACCTGGATGGGTGAATTTAACTCAGATGACCCNNNNNNNNNNNNNNNNNNNNNNNNNN 43220
```

```
695 -----ATTATATCTACTCTCGGCGCAGGAATCCTCAG 728
43221 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 43280
Qy 729 AAGAAATGGAGTAGCCATCATGCTCAACAAAGAGTCCGAAATGCAGTACTTGGATGAG 788
Db 43281 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 43340
Qy 789 TCTCAAAAACGACAGAAATGATCTCTGTTGTTTCCAAAGCGCAAAACCAATCAATATCACAGT 848
Db 43341 TCTCAAAAATGGCAGAAATGATCTCTGTTCAATTCGAAGCAAAACCAATCAATATCACAGT 43400
Qy 849 AATCAAGTCTATGCCCCAACCCAGTAATGCTGAAGAGCTGAAGTTGAACGGTCTCTATGA 908
Db 43401 AATCAAGTCTATGCCCCAACCCAGTAATGCTGAAGAGCTGAAGTTGAACGGTCTCTATGA 43460
Qy 909 AGACCTTACAAGACCTTTTAGAATCAACCCCAAAAGATGCTCTCTCATATATAGGGA 968
Db 43461 AGACCTTACAAGACCTTTTAGAATCAACCCCAAAAGATGCTCTCTCATATATAGGGA 43520
Qy 969 CTGGAATGCAAAAGTAGGAAGCAAAACACCTCGAGTAAACAGGCAAAATTTGGCTTGG 1028
Db 43521 CTGGAATGCAAAAGTAGGAAGTCAAGAAACACCTCGAGTAAACAGGCAAAATTTGGCTTGG 43580
Qy 1029 AATACGGATGAAGCAGGCGCAAGACTAATAGAGTTTCCCAAGAAATGCACTGCTCAT 1088
Db 43581 AATATGGAATGAAGCAGGCGCAAGACTAATAGAGTTTCCCAAGAAATGCACTGCTCAT 43640
Qy 1089 AGCAAAACACCCCTCTCCCAACACACAGAGAGAGCTCTACATGAGACATCAACAGATGG 1148
Db 43641 AGCAAAACACCCCTCTCCCAACACACAGAGAGAGCTCTATACATGAGACATCAACAGATGG 43700
Qy 1149 TCAACACCGAAATCAGATTGATTATTTCTTTGAGCGCAAGATGGAGAGCTCTATACA 1208
Db 43701 TCAACACCGAAATCAGATTGATTATTTCTTTGAGCGCAAGATGGAGAGCTCTATACA 43760
Qy 1209 GTCAGCAAAAACAGACAGGAGCTTACTGTGGCTCAGATCATGAATCTCTTATTTGCCAA 1268
Db 43761 GTCAGCAAAAACAGACAGGAGAACTGACTGTGGCTCAGACCAATGAACTCTTATTTGCCAA 43820
Qy 1269 ATTACAGCTTAAATTGAAGAAAGTAGGGAACCACTAGATCACTCAGTAAAGACTTAA 1328
Db 43821 ATTACAGCTTAAATTGAAGAAAGTAGGGAACCACTAGATCACTCAGTAAAGACTTAA 43880
Qy 1329 TCCATCCCTTATGATTATACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1388
Db 43881 TTAATCCCTTATGATTATACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 43940
Qy 1389 AGACAGAGTACCTAATGAACCTATGGACAGAGGTTTATGACATTTGTACAGGAGCAGGGAT 1448
Db 43941 AGATAGAGTGTCTGATGAACCTATGGATGAGTTCTGTGACATTTGTACAGGAGCAGGGAT 44000
Qy 1449 CGAGACCATCCCATGGAAAGAAATGCAAAAGCAAAATGCGTGTCTGGGGAGGCGCTT 1508
Db 44001 CAAGACCATTTCCATAGAAAGAAATGCAAAAGCAAAATGCGTGTCTGGGGAGGCGCTT 44060
Qy 1509 ACAATATAGTGTGAAAGAGAGAGTGAAGAGCAAAAGCAAAAGCAAAAGCAAAAGCAAAAGCA 1568
Db 44061 ACAATATAGTGTGAAAGAGAGAGTGAAGAGCAAAAGCAAAAGCAAAAGCAAAAGCAAAAGCA 44120
Qy 1569 TCTGAATGAGAGTTCCTCAAA----- 1588
Db 44121 TCTGAATGAGAGTTCCTCAAAAGCAAAAGCAAAAGCAAAAGCAAAAGCAAAAGCAAAAGCAAAAGCA 44180
Qy 1589 ----- 1588
Db 44181 ATGCAGAAATAGAGGAACACACAGAAATGGAAAGACTAGGGATCTCTTCAAGAAAGT 44240
Qy 1589 ----- 1588
Db 44241 CAGAGATACCAAGGAGCAATTTTCATGCAAGAGTGAAGTGAAGAGCAAGAAATGGTAT 44300
Qy 1589 ----- 1588
```


JOURNAL Submitted (24-SEP-2003) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 COMMENT On Sep 24, 2003 this sequence version replaced gi:34328552.
 ----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: <http://www.nisc.nih.gov>
 Contact: nisc.zoo@nhgri.nih.gov
 ----- Project Information
 Center project name: eed
 Center clone name: 234122

QY	840	TATCACAGTAATCCAAAGTCTATGCCCAACCAAGTAATGCTGAAGAAGCTGAAGTTGAAACG	899
DB	36868	TATCACAGTAATCCAAAGTCTATGCCCAACCAAGTAATGCTGAAGAAGCTGAAGTTGAAACG	36927
QY	900	GTCCTATGAGACCTACAAGACCTTTTAGAACCTAAACACCCAAAGAAAGATGCTCTCTCAT	959
DB	36928	GTTCTATGAGACCTACAAGACCTTTTAGAACCTAAACACCCAAAGAAAGATGCTCTCTCAT	36987
QY	960	TATAGGGGACTTGGAAATCCAAAGTAGGAAGCAAGAAACACCTTGGAGTAACAGGCAAAATT	1019
DB	36988	TATAGCAAAATATAATGCAAAAGTAGGAAGTCAAGAAACACCTTGGAGTAACAGGCAAAATT	37047
QY	1020	TGGCTTTGGATACGGAAATGAACAGCGGCAAGACCTAATAGAGTTTGGCCAGAAAGATGC	1079
DB	37048	TGGCTTTGGAAACAGGAAGGAAGCAGGGCAAGACCTAATAGAGTTTGGCCAGAAAGATGC	37107
QY	1080	ACTGGTCATAGCAAAACACCCCTCTTCCAAACAAACAAGAGAAGACTCTACACATGCACATC	1139
DB	37108	ACTGGTCATAGCAAAACACCCCTCTTCCAAACAAACAAGAGAAGACTCTATACATGCACATC	37167
QY	1140	ACCAAGATGGTCAACACCGGAAATCAGATTGATTAATTTCTTTCAGGCCAAAGATGGAGAAG	1199
DB	37168	ACCAAGATGGTCAACACCGGAAATCAGATTGATTAATTTCTTTCAGGCCAAAGATGGAGAAG	37227
QY	1200	CTCTATACAGTCAAGCAAAACCAAGACAGGAGCTTACTGTGGCTCAGATCATGAACCTCT	1259
DB	37228	CTCTATACAGTCAAGCAAAACCAAGACAGGAGCTGACTGTGGCTTAGACCATGAACCTCT	37287
QY	1260	TATTTGCCAAATTCAGACTTAAATTTGAAGAAAGTAGGGAACCACTAGATCACTCAGGTA	1319
DB	37288	TATTTACCAAAATTCAGACTTAAATTTGAAGAAAGTAGGGAACCACTAGACCATTCAGGTA	37347
QY	1320	AGACCTAAATCCCAATCCCTTATCATTTATACAGTGGAAAGTGAAGATAGATTTAAGGGCT	1379
DB	37348	TGACCTAAATCCCAATCCCTTATCATTTATACAGTGGAAAGTGAAGATAGATTTAAGGGCT	37407
QY	1380	AGATCTCATAGACAGAGTACCTAATGAACTATGGACAGAGGTTTCATGACATTTGCAGGA	1439
DB	37408	AGATCTCATAGATAGAGTGCCTCATGAACTATGGAAATGAGTTCTGTGACACAGTACAGGA	37467
QY	1440	GACAGGATCGAGACCATCCCAATGAAAAGAAATGCAAAAAGCAAAATGGCTGTCTGG	1499
DB	37468	GACAGGATCAAGACCATCCCAATGAAAAGAAATGCAAAAATC-AAATGGCTATCTGG	37526
QY	1500	GGAGGCCTTACAAATAGCTGTGAAAAAGAGAAGTGAAGCAAGCAAGAAAGGAAAG	1559
DB	37527	AGAGGCCTTACAAATAATTGTGAAGAGAGAGAGCAAAAGCAAGCAAGAAAGAGAG	37586
QY	1560	ATAAAGCATCTGAATGCAGAGTTCCAAA-	1588
DB	37587	ATATAACATCTGAATGTAGAGTTCCAAAAGAAATAGCAAGAGAGATAGAAGGCTTCTT	37646
QY	1589	-----	1588
DB	37647	CAGTGATCAATCAAGAAATAGAGCAAAACAAACAGATGGGAAGACTACCAATCTCTT	37706
QY	1589	-----	1588
DB	37707	CAAGAAAAATCAGCGATAGCAAGGAACATTTTCATGCAAAAGATGAGCTCTATAAAGACAG	37766
QY	1589	-----	1588
DB	37767	AAAGGTATGGACCTAACAGAGCAGACAGATATTAAAGAGAGATGGCAAGAAATACAGGA	37826
QY	1589	-----	1588
DB	37827	AGAACTGTACAAAAAGATCTTTCATGACCAAGATATACGATGGTGTGTGATCACTGACCTA	37886
QY	1589	-----	1588
DB	37887	GAGCCAGACATCTCGAATGTGAAGTCAAGTGGGCTTCAAGAGCATCACTACAAACAAA	37946
QY	1589	-----	1588

37947	DB	GCTAGTGGAGGTGATGGAAATTCAGTTTGAGCTATTCCAAATCCTGAAAGATGATGCTGTG	38006
1589	QY	-----	1588
38007	DB	AAAGTGTGCACCTCAATATGCCAGCAAAATTTGGAAAACTCAGCAATGCCACAGGACTGG	38066
1589	QY	-----	1588
38067	DB	AAAAGGCGAGTTTTCATTCCAATCCAAAGAAAGGCAATGCCAAAGAATGCTCAAACCTAC	38126
1589	QY	-----	1588
38127	DB	CGCACATTCGACTCATCTCACAGCTAGTAAAGGAATGCTCAAATTTCTCCAAGTCAGG	38186
1589	QY	-----GAACTTCAGTTGTTCAAGCTGGTTTTAGAAAACTCAGA	1627
38187	DB	CTTCAGCAATACGTTGAACCATGAACCTCTGATGTTCAAGCTGGTTTTGCGAAAAGCGAGA	38246
1628	QY	GGAAACAGACACAAATTTGCAACATCCTCTGATCATGGAAGGAGCAAGAGAGTTCCAG	1687
38247	DB	GGAAACAGAGCCCAATTTGCCAACATCTCTCGATCATAGAAAAAGCAAGAGAGTTCCAG	38306
1688	QY	AAAAACATCTATTTCTGCTTTTATGTGCTATGCAAAAGCCTTTTGACTGTGGGGGTCACAAT	1747
38307	DB	AAAAACATCTATTTCTGCTTTTATGTGCTATGCAAAAGCCTTTTGACTGTGGATCACTAT	38366
1748	QY	AAACTGTGGAAAAATTCGAAAGGATGGGAATACAGAGCCACTCACCTGACTCTTTGAAA	1807
38367	DB	AAACTGTGGAAAAATTCGAAAGGATGGGAATACAGAGCCACTCACCTGACTCTTTGAGA	38426
1808	QY	AATTTGATGCAAGTCAGGAAGCAACGTTAGAACTGGACATGGAAACACAGACTGGTTC	1867
38427	DB	AACTGTATGCAGGTCAGGAAGCAACGTTAGAACTGGACATGGAAACACAGACTGGTTA	38486
1868	QY	CAAGTAGGAAAAAGGATGTCAGAGGCTGTATTTGTCCACCGGCTGTTTAACTTCTAT	1927
38487	DB	CAATAGGAAAAAGGATGTCAGAGGCTGTATTTGTCCACCGGCTGTTTAACTTCTAT	38546
1928	QY	GCAGAG-ACATCATGAGAAACCTCGGCTGGAAGAGCACAGCTGGAATCAAGATTGCC	1986
38547	DB	GCAGAGTACATCATGAGAAACCTCGACTGGAAGAAACACAAGCTGGAATCAAGATTGCC	38606
1987	QY	GGGAGAAATAGCAATTAACCTCAGATATGCAGA-TGATACCAACCTTATGGCAGAAAGTGA	2045
38607	DB	AGAGAAATCTCAATTAACCTCAGATATGCAGATTGACACCCCTTATGGCAGAAAGTGA	38666
2046	QY	AGAGGAACATAAAGCCTCTTGATGAAGGTGAAGAGGAGAGTGAAAGGTGGCTTTAAA	2105
38667	DB	AGAGGAACATAAAGCCTCTTGATGAAGGTGAAGGTGGAGAGTGAAAGGTGACTTTAAA	38726
2106	QY	GCTCAACATTCAGAAAGCAAGATCATGGCATCTGTCCTCATCTTCTATGGGAAATAGA	2165
38727	DB	GCTCAACATTCAGAAAC---GATCATGGCATCCAGTCCCATCACTTCATGGGAAATAGA	38783
2166	QY	TGGGGAAAACAGTGGAAAACAGTGTCCAGATTATTTTGGGGGGGCTCCAAATCACTGCCA	2225
38784	DB	TGGGGAAAACAGTGGAAAACAGTGTCCAGATTATTTT---CGGGGCTCCAAATCACTACA	38841
2226	QY	GATGGTGACTGCAGCCATGAATTAAGAACACATTACTCTTTGGAAGAAAGTTAACCAA	2285
38842	DB	GATGGTGACTGCAGCCATGAATTA---TGACCAA	38874
2286	QY	CTTAGATAGCATATTGAAAAGC--AGAGACATTACTTTGCCAAACAAAGCCCATCTAGTC	2343
38875	DB	CATAGATACATATTCAAAGCAAGAGACATTACTTTGTCAACAAAGGTTCACTTAGTC	38934
2344	QY	AAGGCTATGGTTTTCCAGTGGTCATGTATGGATGTGAGAGTTGGACTGTGAAGAAAGCT	2403
38935	DB	AAGGCTATGGTTTTCCAGTGGTCATGTATGGATGTGAGAGTTGGACTGTGAAGAAAGGCT	38994
2404	QY	GAGCACTGAAGAAATTCATGCTCTTTTGAACCTGTGGTGTGGAGAAAGACTCTTGAGAGTCCCT	2463

Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoeningri.nih.gov
----- Project Information
Center project name: crk
Center clone name: 056C13

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics -----
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; version 0.990319
Assembly program: Phrap; version 0.990319
Consensus quality: 123475 bases at least Q40
Consensus quality: 123571 bases at least Q30
Consensus quality: 123601 bases at least Q20
Insert size: 111000; agarose-fp
Quality coverage: 10.67x in Q20 bases; agarose-fp
Quality coverage: 9.58x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

* 1 35692: contig of 35692 bp in length
* 35693 123710: contig of 87918 bp in length.

Location/Qualifiers
1. 123710
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="RP42-56C13"
/clone_lib="RP42"
/note="assembly_fragment
clone end:T7
vector side:left"
35793_123710
/note="assembly_fragment
clone end:SP6
vector side:right"
11382_123710
/note="clone overlaps with GenBank Accession Number
AC116433 clone RP42-556E7 (center project name crk)"

Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoeningri.nih.gov
----- Project Information
Center project name: crk
Center clone name: 056C13

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics -----
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; version 0.990319
Assembly program: Phrap; version 0.990319
Consensus quality: 123475 bases at least Q40
Consensus quality: 123571 bases at least Q30
Consensus quality: 123601 bases at least Q20
Insert size: 111000; agarose-fp
Quality coverage: 10.67x in Q20 bases; agarose-fp
Quality coverage: 9.58x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

* 1 35692: contig of 35692 bp in length
* 35693 123710: contig of 87918 bp in length.

Location/Qualifiers
1. 123710
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="RP42-56C13"
/clone_lib="RP42"
/note="assembly_fragment
clone end:T7
vector side:left"
35793_123710
/note="assembly_fragment
clone end:SP6
vector side:right"
11382_123710
/note="clone overlaps with GenBank Accession Number
AC116433 clone RP42-556E7 (center project name crk)"

Query Match 57.4%; Score 2048.6; DB 1; Length 123710;
Best Local Similarity 76.2%; Pred. No. 0;
Matches 2592; Conservative 0; Mismatches 194; Indels 614; Gaps 12;
QY 1 GTCAGGAAGGGCGGCGAGTGGAGAGGTACCTCGTCCAAAGTAAGGACGACTAGC 60
DB 109448 GTGAGGAGGGCGGTGCTGAGGAGA-----TACCCCTCTCCAAAGGAGGACGCTGGC 109502
QY 61 TGCCTTTTGTGAGGAGCGCGTAAAGATACCCGCGCCCAAGGTAAAGAAACCCAG 120
|||||

Db 109503 TGTGCTTTGCTGGAGCAGTGTGTGAAGAGATATCCCATGTCTAAAGTAAAGAAACCCCAAG 109562
Qy 121 TAAGATGTTAGTGTGTTGAGAGGCATCAGAGGCGACACATATCTGAAACCATACAGCA 180
Db 109563 TAAGATGTTAGTGTGTTGAGAGGCATCAGAGGCGACACATGAAACCATATATCAC- 109621
Qy 181 GAAAACTAGTCAATCAATCACTACACTAGGACACAGCCCTTGTCTAACTCAATGAAACTAAG 240
Db 109622 GAAAACTAGTCAATCAATCACTACACTAGGACACAGCCCTTGTCTAACTCAATGAAACTAAG 109679
Qy 241 COATCCCC-GTGGGCAACCAAGATGGCGAGGTGATGCTGGAGAGATCTGACAGATGT 299
Db 109680 CCATGCCCTGTGGGCGCCACCAAGACGACAGGTGATGCTGGAGAGGTCTGACAGATGT 109739
Qy 300 GGTCCACTGGAGAGGGGAATGCAACCACTTCAGTATTCCTTGGCTTGGAAACCCCATGAA 359
Db 109740 GGTCCACTGGAGAGGGGAATGCAACCACTTCAGTATTCCTTGGCTTGGAAACCCCATGAA 109799
Qy 360 CAGTATGAAAGGCAAAATGATAGATACCTGAAAGAGGAATCTCCCAAGGTGAGTAGTGC 419
Db 109800 CAGTATGAAAGGCAAAATGATAGATACCTGAAAGAGGAATCTCCCAAGGTGAGTAGTGC 109859
Qy 420 CCCATATCTCTGAGATCAGTGGAGAAATAAATCCAGAAAGAAATGAAGAGATGGAGCC 479
Db 109860 CCAATACCTCTAGTGGAGATCAGTGGAGAAATAAATCCAGAAAGAAATGAAGAGATGGAGCC 109919
Qy 480 AAAGCAAAAGAAATACCCAGCTGGATGTGACTGGTGTATATAAGCAAGTCCGATGCTG 539
Db 109920 AAAACAAAACCAATACCCAGTGTGGATGTGACTGGTGTATATAAGCAAGTCCGATGCTG 109979
Qy 540 TAAAGACAATATTCATAGAACCTGGAATGCTGAGTCCATGAATCAAGCCAAATTTGA 599
Db 109980 TAAAGACAATATTCATAGAACCTGGAATGCTGAGTCCATGAATCAAGCCAAATTTGA 110039
Qy 600 AGTGTCAAACAAGAGATGGCAAGAGTGAATGTCAACATCTAGGAATCAGCAAACTAAA 659
Db 110040 AGTGTCAAACAAGAGATGGCAAGAGTGAATGTCAACATCTAGGAATCAGCAAACTAAA 110099
Qy 660 ATGGACTGGAATGGTGAATTTAATCTAGATGACATTTATCTACTGCGGGCAGGA 719
Db 110100 ATGGACTGGAATGGTGAATTTAATCTAGATGACATTTATCTACTGCGGGCAGGA 110159
Qy 720 ATCCCTCAGAGAAATGAGTAGCCATCATGCTCAACAAAGAGTCCGAAATGCACTACT 779
Db 110160 ATCCCTTAGAGAAATGAGTAGCCATCAT-GTCAACAAAGAGTCTGAAATGCACTACT 110218
Qy 780 TGGATGCACTCAAACACGACAGATGATCTGTGTGTTTCCAGGCAAAACCAATCAA 839
Db 110219 TGGCTGCAATCTCAAAAAGAACAGAAATATCTGTGTTCTTCCAGGCAAAACCAATCAA 110278
Qy 840 TATCAGATATCAAGTCTATGCCCAACAGTAATGCTGAAAGAGTCAAGTGAACG 899
Db 110279 TAGCAGATATCAAGTCTATGCTCCAACAGTAATGCTGAAAGAGTCAAGTGAACG 110338
Qy 900 GTCCTATGAAGACCTACAAGACCTTTTAGAATCAACACCAAAAGAGATGCTCCTCTCAT 959
Db 110339 GTTCTATGAAGACCTACAAGACCTTTTAGAATCAACACCAAAAGAGATGCTCCTCTCAT 110398
Qy 960 TATAGGGGACTGGAATGCAAAAGTAGGAAGCAAGAAACACCTGGAGTACAGGCAAAAT 1019
Db 110399 TATAGGGGACTGGAATGCAAAAGTAGGAAGTCAAGAAACACCTGGGTAACAGGCAAAAT 110458
Qy 1020 TGGCTTTGGAAATACGGAATGAAGCAGGCGCAAGAACTAATAGATTTTGGCAAGAAATGC 1079
Db 110459 TGGCTTTGGAAATGAAGCAGGCGCAAGGCTAATAGATTTTGGCAAGAAATGC 110518
Qy 1080 ACTGGTATAGCAAAACCCCTCTTCCAAACACCAAGAGAGACTCTACATGAGCATC 1139
Db 110519 ACTGGTATAGCAAAACCCCTCTTCCAAACACCAAGAGAGACTCTACATGAGCATC 110578
Qy 1140 ACCAGATGGTCAACACCGAAATCAGATTGATTATTTCTTTGACGCCAAAGATGGAGAG 1199
Db 110579 ACCAGATGGTCAACACCGAAATCAGATTGATTATTTCTTTTGGCAGCAAGATGGAGAG 110638

Qy 1200 CTCTATACAGTCAGCAAAACCAAGACCAGAGCTTACTGTGGCTCAGATCATGAATCTCT 1259
Db 110639 CTCTATACAGTCAGCAAAACCAAGCCCAGGAGCTACTGTGGCTCAGATCATGAATCTCT 110698
Qy 1260 TATTGCAAAATTCAGACTTTAAATTTGAAGAAAGTAGGGAACCACTAGATCACTCAGGTA 1319
Db 110699 TATTGCAAAATTCAGACTTTAAATTTGAAGAAAGTAGGGAACCACTAGATCACTCAGGTA 110758
Qy 1320 AGACCTAAATCCAAATCCCTTATGATTTATACGTGGAAGTGAAGATAGATTTAAAGGCT 1379
Db 110759 TGACCTAAATCAAAATCCCTTATGATTTATACGTGGAAGTGAAGATAGATTTAAAGGCT 110818
Qy 1380 AGATCTGATAGACAGAGTACCTAAATGAACATATGGACAGAGGTTCTGACATTTGACAGGA 1439
Db 110819 AGATCTGATAGACAGAGTACCTGATGAACCTGTGGACAGAGTTCTATACATTTGACAGGA 110878
Qy 1440 GACAGGATCGAGACCATCCCATGGAAGAAATGCAAAAGCAAAATGGCTGTCTGG 1499
Db 110879 GACAGGATCGAGACCATCCCATGGAAGAAATGCAAAAGCAAAATGGCTGTCTGG 110938
Qy 1500 GGAGGCTTTACAAATAGCTGTGAAAGAGAGAGTGAAGAAAGCAAGGAAAGGAAAG 1559
Db 110939 GGAGGCTTTACAAATAGCTGTGAAAGAGAGAGTGAAGAAAGCAAGGAAAGGAAAG 110998
Qy 1560 ATAAAGCATCTGAATGACAGATTTCCAAA- 1588
Db 110999 ATACAAACCTCTGAATGCAGAGATCCAAAGAAATAGAAAGGAGAGATAGAAAGCCTTTCT 111058
Qy 1589 - 1588
Db 111059 CAGGATCAATCAAAAGAAATAGAGGAAACAAACAGATGGGAAGACTAGAAATCTCTT 111118
Qy 1589 - 1588
Db 111119 CAAGACAATTAGAGATACCAAGGAATATTTTCATGCAAGATGGCTCGATAAGGACAG 111178
Qy 1589 - 1588
Db 111179 AAAAGGTATGATCTAACAGAGCAGAAAGATATTAAAGAGAGGTGGCAAGAAATACAGGA 111238
Qy 1589 - 1588
Db 111239 AGAACTGTATAAAAGATCTTCATGACCAAGATATCACAATGATGTAACTCACTCACT 111298
Qy 1589 - 1588
Db 111299 AGAGCCAGACATCTCTGGAATGTGAAGTCAAGTGGGCTTTAGAAAGCATCACTACGAAACA 111358
Qy 1589 - 1588
Db 111359 AGCTAGTGAGATGATGGAATTCAGTTAAGCTATTTCAAAATCTGAAAGATGATGCTGT 111418
Qy 1589 - 1588
Db 111419 GAAAGTGTGCACTCAATATGCCAAATTTGGAAACCGCAGCAGTGGCCACAGGACTG 111478
Qy 1589 - 1588
Db 111479 GAAAGGTGAGTTTTCATTTCCATTCCTCAATGCCAAAGAGGAAATGCCAAAGATGCTCAAACTA 111538
Qy 1589 - 1588
Db 111539 CCTCACAATTTGACTCATCTCACAGCTAGTAAAGTAATGCTCAAAATTTCTCAAGCCAG 111598
Qy 1589 - 1588
Db 111599 GCTTCAGCAATACATGAACCGTGAACCTTCCAGATGTTCAAGCTGTTTGAAGAAAGCAG 1626
Qy 1627 AGGAACCAAGACCAAAATTTGCAACATCTCTGATCATGAAAGGAAAGCAAGAGAGTTCCTCA 1686
Db 111659 AGGAACCAAGATCAAAATTTGCAACATCTCTGATCTTTGAAAGGCAAGAGATTCCTCA 111718

Qy	1587	GA AAAA CATCTATTTCTGCTTTATTTAGCTATGCAAAAGCCCTTTGACTGTGGGGTCAAA	1746
Db	111719	GA AAAA CATCTATTTCTGCTTTATTTAGCTATGCAAAAGCCCTTTGACTGTGGGTACAA	111778
Qy	1747	TAAACTGTGGA AAAATTTCTGAAAGGATGGGAATACCAGACCACTTGAC- --CTGACTCTT	1803
Db	111779	TAAACTGTGGA AAAATTTCTGAAAGAGATGGGAATACCAGACCACTTGACTTCTTCTCTCT	111838
Qy	1804	GA AAAA TTTGTATGCAGGTCAAGAGCAACAGTTAGAACTGCACTGCAACACAGACTG	1863
Db	111839	GAGAAA CCTATATGCAGGTCAAGAGCAACAGTTAGAACTGCACTGCAACACAGACTG	111898
Qy	1864	GTTCCAAAGTAGAAAAGAGGATGTCAAGGCTGTATATTGTCAACCGGCTTGTTTAACTT	1923
Db	111899	GTTCCAAATAGAAAAGAGGTACATCAAGGCTGTGTATTGTCAACCTGTATTATTAACTT	111958
Qy	1924	CTATGCAAG-ACATCATGAAAACGCTGGGCTGGAAGACAAAGTGGAACTGGAATCAAGAT	1982
Db	111959	CTATGCAAGTACATCATGAAAACACCTGGGCTGAAAAGAACGAGCTGGAATCAAGAT	112018
Qy	1983	TGCCGGAGGAATAGCAATACCTCAGATATGCAAGTATACCACTTATGGCAGAAAG	2042
Db	112019	TGCCGGAGGAATATCAATACTCAGATATACAGTACACCACTTATGGCAGAAAG	112078
Qy	2043	TGAAGAGGAACATAAAAACCTCTTGATGAAGGTGAAAGAGGAGTGA AAAAGTTGGCTT	2102
Db	112079	TGAAGAGGAACATAAAAACCTCTTGATGAAGGTGAAAGAGGAGTGA AAAAGTTGGCTT	112138
Qy	2103	AAAGCTCAACATTCAGAAAAAGAGATCATGGCATCTGGTCCCACACTTCATGGGAAT	2162
Db	112139	AAAGCTCAACATTCAGAAATCTAAGATCATGGCATCTGGTCCCACACTTCATGGGAAT	112198
Qy	2163	AGATGGGAAAACAGTGGAAAACAGTGTACAGCTTTATTTTTGGGGGGCTCCAAAATCACT	2222
Db	112199	AGATGGGAAAACAGTGC AAAACAGTGTCCGACTTTA- -TTTTGGGGGGCTACAAAATCACT	112256
Qy	2223	GCA- -GATGTGACTGCAGCCATGAAAATTA AAAGACACTTACTCTCTTGGAGAAAAGTTA	2280
Db	112257	GCAATCACTGTGATTGCAGCCATGAAAATTA AAAGATACTTACTCTCTTGGAGAAAATTA	112316
Qy	2281	--ACCAACTAGATAGCATATTTGAAAACGAGACATACCTTGC CAACAAAGCCCCATC	2338
Db	112317	TGACCAACTAGATAGCATATTTAAAAGCAGACACTATTTTGGCAAAAGTCTGTC	112376
Qy	2339	TAGTC- AAGGCTATGGTTTTTCCAGTGTGATGTATGGATGTGAGATTGGA CTGTGAAG	2397
Db	112377	TAGTCAAAAGGCTATGGTTTTTCTGTGTGATGTATGGATGTGAGATTGGA CTGTGAAG	112436
Qy	2398	AAAGCTGAGCACTGAAGAAATTGATGCTTTTGAAGTGTGGTGTGAGAGAGACTCTTGAG	2457
Db	112437	AAAGCTGAGTGTGAGAAATTTGATGCTTTTGAAGTGTGGTGTGAGAGAGACTCTTGAG	112496
Qy	2458	GTCCCTCTGCACTGCAAGGAGATCCAACAGATCCATTCTGAAGGAGATCAGCCCTGGGATT	2517
Db	112497	GTCCCTCTGAACTGCAAGGAGATCCAACCGTCCATTCTAAAGGAGATCAGTCTGGGTGT	112556
Qy	2518	TCTTTTGAAGGAATGATGCTAAAGCTGAAACTCCAGTACTTTTGGCCACTGTATCAGAGA	2577
Db	112557	TCAITTTGAAGGACTGATGCTAAAGCTGAAATTTCCAGTACTTTTGGCCACTCTGCAAGA	112616
Qy	2578	GCTGACTCACTGAAAAGACCTGTATGCTGGAGGGATTGGGGCAGAGGAGAGAGGGGA	2637
Db	112617	GTTGACTCACTGAAAAGACTCTGATGCTGGAGGGATTGGGGCAGAGGAGAGAGGGGA	112676
Qy	2638	CGACAGAGGATGAGATGGCTGGATGGCATCACTGACTGATGGA CGTGAGTCTGGGTGAA	2697
Db	112677	TGACAGAGGATGAGATGGCTGGATGGCATCACTGACTGATGATGAGTGTGAGTGAAT	112736
Qy	2698	CTCCTGGAGTTGGTCACTGACAGGGAGGCTCTGCTGGCGGATTCATGGGCTCAAAAG	2757
Db	112737	CTCTGGGAGTTGGTCACTGACAGGGAGGCTCTGGCGTCTGCCGATTCATGGGCTGCCAAG	112796
Qy	2758	AGTTGGACACGACTGAGCAACCTGAACTGAACTGAACTGTA 2797	

Db 112797 AGTCGGACATGACTAGTGACTGAAGTAAGTGAA 112836

RESULT 26
AC091399 linear HTG 01-JUN-2002
LOCUS Bos taurus clone RP42-65K24, WORKING DRAFT SEQUENCE, 5 ordered pieces
DEFINITION AC091399 181921 bp DNA
ACCESSION AC091399 GI:2130657
VERSION HTG; HTGS_PHASE2; HTGS_DRAFT.
KEYWORDS Bos taurus (cow)
SOURCE Bos taurus
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 181921)
AUTHORS Akhtar,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E., Laric,P., Lee-lin,S.-O., Legaspi,R., Maduro,O.L., Maduro,V.B., Margulies,E.H., Masilelo,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Paquirigan,C., Pearson,R., Portnoy,M.E., Prasad,A., Schueler,M.G., Stancirpop,S., Thomas,J.W., Thomas,P.J., Touchman,L.W., Tsugeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,J., Young,A., Zhang,L.-H. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 181921)
Green,E.D.
Direct Submission
Submitted (19-APR-2001) NIH Intramural Sequencing Center, 8717 Government Circle, Gaithersburg, MD 20877, USA
JOURNAL
REFERENCE 3 (bases 1 to 181921)
Green,E.D.
Direct Submission
Submitted (01-JUN-2002) NIH Intramural Sequencing Center, 8717 Government Circle, Gaithersburg, MD 20877, USA
JOURNAL
COMMENT On Jun 1, 2002 this sequence version replaced gi:13677070.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoon@nigr.nih.gov
----- Project Information
Center project name: axb
Center clone name: 065K24

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

```

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 180679 bases at least Q40
Consensus quality: 181147 bases at least Q30
Consensus quality: 181147 bases at least Q20
Consensus quality: 181403 bases at least Q20
Insert size: 178000; agarose-ftp
Insert size: 181521; sum-of-contigs
Quality coverage: 1.136x in Q20 bases; agarose-ftp
Quality coverage: 1.14x in Q20 bases; sum-of-contigs

```

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 5 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * been provided by the submitter.

* This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

* 1 27632: contig of 27632 bp in length
 * 27633 27732: gap of unknown length
 * 27733 69608: contig of 41876 bp in length
 * 69609 69708: gap of unknown length
 * 69709 109185: contig of 39477 bp in length
 * 109186 109285: gap of unknown length
 * 109286 181783: contig of 72498 bp in length
 * 181784 181883: gap of unknown length
 * 181884 181921: contig of 38 bp in length.

FEATURES

source

1.181921

/organism="Bos taurus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9913"
 /clone="RP42-65K24"
 /clone_lib="RP42"

1..75601

/note="clone overlaps with GenBank Accession Number
 AC091660 clone RP42-354B6 (center project name axa)"

1..27632

/note="assembly_fragment"

clone_end:SP6

vector_side:left

27733..69608

/note="assembly_fragment"

69709..109185

/note="assembly_fragment"

109286..181783

/note="assembly_fragment"

181884..181921

/note="assembly_fragment"

clone_end:T7

vector_side:right

Query Match 57.3%; Score 2046.4; DB 1; Length 181921;
 Best Local Similarity 74.1%; Pred No. 0;
 Matches 2592; Conservative 0; Mismatches 201; Indels 706; Gaps 9;

QY	1	GTCCAGGAGGCGGCGAGTGAGAGAGTACCTCCCTGCTCCAAAGGTAAAGGAGCAGTACG	60
DB	66540	GTCCAGGAGGCGGCGGTGAGGAGA-----CACCCCTCGTCCAAAGGTAAAGGAGCAGTACG	66594
QY	61	TGCCCTTGCTGGAGCAGCCGTAAGAGATACCCAGCCCGCCAGGTAAAGAGAACCCAG	120
DB	66595	TGCCCTTGCTGGAGCAGCCGTAAGAGATACCCAGCCCGCCAGGTAAAGAGAACCCAG	66654
QY	121	TAAGATGGTAGGTGTGTGAGAGGGGCATCAGAGGGCGAGACATCTGAAACCATACAGCA	180
DB	66655	TAAGAGCGTAGGTGTGCAAGAGGGGCATCAGAGGGCGAGACATCTGAAACCATATCAT	66714
QY	181	GAAACTAGTCAATCTAATCACTAGGACAGACAGCTTGTCTCAATCAATGAACTAAG	240
DB	66715	GAAACTAGTCAATCTAATCACTAGGACAGACAGCTTGTCTCAATCAATGAACTAAG	66774
QY	241	CCATGCCCGTGGGCAACCCCAAGATGGCAGGTCAATGGTGAGAGATCTGCACAGAATGTG	300
DB	66775	CCATGCCCGTGGGCAACCCCAAGATGGCAGGTCAATGGTGAGAGATCTGCACAGAATGTG	66834
QY	301	GTCCACTGGAGAGGGGAAT-GCAAAACACATTCAGTATTTCTTGCCTTGAGAACCCCATGAA	359
DB	66835	GTCCACTGGAGAGGGGAATGCAAAACACATTCAGTATTTCTTGCCTTGAGAACCCCATGAA	66894
QY	360	CAGTATGAAAAGGCAAAATGATAGGATCTGAAAGAGGAACTCCCGCAGGTCAGTAGGTGC	419

DB	66895	CAGTATGAAAAGGCAAAATGATAGGATACAGAAAGAGGAATTCGCCAGGTCAGTAGGTGC	66954
QY	420	CCCATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGAAATGAAGAGATGAGCC	479
DB	66955	CCAGTATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGAAATGAAGAGATGAGCC	67014
QY	480	AAAGCAAAAAGAAATACCCAGCTGTGGATGTGATGTGATGTGATGTGATGTGATGTGATGTG	539
DB	67015	AAAGCAAAAAGAAATACCCAGCTGTGGATGTGATGTGATGTGATGTGATGTGATGTGATGTG	67074
QY	540	TAAAGAGCAATATTCATAGGAACCTGGAAATGTCAGGTCCATGAATCAAGGCAAAATGGA	599
DB	67075	TAAAGAGCAATATTCATAGGAACCTGGAAATGTCAGGTCCATGAATCAAGGCAAAATGGA	67134
QY	600	AGTGGTCAAAACAGAGATGCGCAAGAGTGAATGTCAACATCTTAGGAATCAGGCAAACTAAA	659
DB	67135	AGTGGTCAAAACAGAGATGCGCAAGAGTGAATGTCAACATCTTAGGAATCAGGCAAACTAAA	67194
QY	660	ATGGACTGGAATGGGTGAATTTAACTCAGATGACCATTTATATCTACTACTTGGCGGACGA	719
DB	67195	ATGGACTGGAATGGGTGAATTTAACTCAGATGACCATTTATATCTACTACTTGGCGGACGA	67254
QY	720	ATCCCTCAGAAGAAATGGAGTAGCCATCATGTCAACAAAGAGTCCGAAATGCGAGTACT	779
DB	67255	ATCCCTCAGAAGAAATGGAGTAGCCCTTGTACAGAAACAAAGAGTCCGAAATGCGAGTACT	67314
QY	780	TGATGCTAGTCTCAAAAACGACAGAAATGATCTCTGTTTTCAGGCAAAACCAATCAA	839
DB	67315	TGATGCTAGTCTCAAAAACGACAGAAATGATCTCTGTTTTCAGGCAAAACCAATCAA	67374
QY	840	TATCACAGTAATCCAGTCTATGCCCCACAGTAACTGCTGAAGAGCTGAAGTTGAACG	899
DB	67375	TATCACAGTAATCCAGTCTATGCCCCACAGTAACTGCTGAAGAGCTGAAGTTGAACG	67434
QY	900	GTCTTATGAAGACCTACAAGACCTTTTATAGAACTAAACCCCAAAAGAGATGCTCTTCTCAT	959
DB	67435	GTCTTATGAAGACCTACAAGACCTTTTATAGAACTAAACCCCAAAAGAGATGCTCTTCTCAT	67494
QY	960	TATAGGGAGTGAATGCAAAAGTAGGAAGCAAGAAACACCTGAGTAACAGGCAAAAT	1019
DB	67495	TATAGGGAGTGAATGCAAAAGTAGGAAGTCAAGAAACACCTGAGTAACAGGCAAAAT	67554
QY	1020	TGCGCTTGGATACGGAATGAAGCAGGCAAGACATAATAGAGTTTTCGCAAGAAATATGC	1079
DB	67555	TGCGCTTGGATACGGAATGAAGCAGGCAAGACATAATAGAGTTTTCGCAAGAAATATGC	67614
QY	1080	ACTGGTATAGCAAAACACCTCTTCCAAACACACAGAGAGAGACTCTACATATGGACATC	1139
DB	67615	ACTGGTATAGCAAAACACCTCTTCCAAACACACAGAGAGAGACTCTACATATGGACATC	67674
QY	1140	ACCATGCTGTAACCCGAAATCAGATTGATTATATCTTTTCGACCCAAAGATGAGAGAG	1199
DB	67675	ACCATGCTGTAACCCGAAATCAGATTGATTATATCTTTTCGACCCAAAGATGAGAGAG	67734
QY	1200	CTCTATACAGTACAGCAAAACCAAGACCCAGGAGCTTACTGTGGCTCAGATCATGAACTCCT	1259
DB	67735	CTCTATACAGTACAGCAAAACCAAGACCCAGGAGCTTACTGTGGCTCAGATCATGAACTCCT	67794
QY	1260	TATTCGCAAAATTCAGACTTAATTCGAAGAAAGTAGGGAACCACTGATCAGTACAGGTA	1319
DB	67795	TATTCGCAAAATTCAGACTTAATTCGAAGAAAGTAGGGAACCACTGATCAGTACAGGTA	67854
QY	1320	AGACCTTAATCAATCCCTTATGATTATATACAGTGAAGTGAAGAAATAGATTAAAGGGCCT	1379
DB	67855	AGACCTTAATCAATCCCTTATGATTATATACAGTGAAGTGAAGAAATAGATTAAAGGGCCT	67914
QY	1380	AGATCTGATAGACAGAGTACCTTAATGAAGTATGACAGAGGTTTATGACATGTACAGGA	1439
DB	67915	AGATCTGATAGACAGAGTACCTTAATGAAGTATGACAGAGGTTTATGACATGTACAGGA	67974
QY	1440	GACAGGATACGAGACCATCCCATCGAAAAAGAAATGCAAAAAAGCAAAATGGCTGTCTGG	1499

Db 67975 GACAGGATCAAGACCATCTCTTTGGAAAAAGAAATGCAAAAAAGCAAAATGGCTGTCTGG 68034
Qy 1500 GGAGCCCTTACAAATAGCTGTGAAAGAGAGAGAGTGAAGAGCAAGGAAAAAGGAAAG 1559
Db 68035 GGAGCCCTTACAAATAGCTGTGAAAGAGAGAGAGTGAAGAGCAAGGAAAAAGGAAAG 68094
Qy 1560 ATAAAGCATCTGAATGCAGAGTTCCAAA 1588
Db 68095 ATATAAGCATCTGAATGCAGAGTTCCAAAAGATAGCAAGAGAGATGAAGAAAGCCTTCCT 68154
Qy 1589 1588
Db 68155 CAGTGATCAATGCAAGAAATAGAGGAAATATACAGAAATGGGAAAGACTAGAGATCTCTT 68214
Qy 1589 1588
Db 68215 CAAGAAATATAGATACCAAGGGAACATTTTCATGCAAGATGGCTCGATAAAGGACAG 68274
Qy 1589 1588
Db 68275 AAATGGTATGCCTTAACAGGAGCAGAGATATTTAAAAAGAGGTGGCAGGAATACAGA 68334
Qy 1589 1588
Db 68335 AGAATGTACAAAAAGAGCTTCAGACCCAGATAATCAAGATGGTGTGATCACTGACCT 68394
Qy 1589 1588
Db 68395 AGAGCCAGACATCTCGAATGTGAAGTCAAGTGGGCTTAGAAAAGCATCTCTACAAAAA 68454
Qy 1589 1588
Db 68455 AAAAAAAGAAAGCATCTTACGAAACAAAGTGTGAGGTGATGAAATTCAGTGGAG 68514
Qy 1589 1588
Db 68515 CTATTTCAAATCCTCAAGATGATGCTGTGAAAGTGTGCATCAATATGCCAGAAAT 68574
Qy 1589 1588
Db 68575 TGGAAACTCAGAGTGGCCACAGAGCTGGAAGAGGTCAAGTTTCATTCCTCAATCCCAAA 68634
Qy 1589 1588
Db 68635 AAAGCAATGCCAAGATGCTCAAACTACTGACAAATTTGACTCAATCTCAGATGTAGT 68694
Qy 1589 1588
Db 68695 AAAGTAATGCTCAAAATTTCTCAAGTAGGCTTCAGCAATATGTCAACCGTGAATCTCT 68754
Qy 1598 GTTGTTCAGCTGGTTTGAAGAACTCAGAGGAAACAGAGACCAAAATTTGCCAACATCTC 1657
Db 68755 GATGTTCAAGCTGGTTTGAAGAAAGGAGGAAACAGAGATCAAAATTTGCCAATATCCAC 68814
Qy 1658 TGATCATGGAAGAAAGCAGAGAGTTCCAGAAAAACATCTATTTCTGCTTTATTGACTAT 1717
Db 68815 TGGATCATCAAAAAGCAGAGAGTTCCAGAAAAACATCTATTTCTGCTTTATTGACTAT 68874
Qy 1718 GCAAGAGCTTTGACTGTGGGGTCAAACTGTTGGAATTTCTGAAAGGATGGA 1777
Db 68875 GCAAGAGCTTTGACTGTGTGATCAAAATGAAGTGTGAAATTTCTGAAAGAGATGGA 68934
Qy 1778 ATACCAAGCACCCTGACCTGACTCTTTGAAAAATTTGTATGCAAGTTCAGGAGTCAAGCAACAGTT 1837
Db 68935 ATACCAAGCACCCTGACCTGCTCTTTGAAACCTATATACAGTTCAGGAGCAACAGTT 68994
Qy 1838 AGAATGACATGGAACACAGACTGGTTCCAGTAGGAAAGAGTATGTCAAGCTGT 1897
Db 68995 AGAATGCGCATGGAACACAGACTGGTTCCAAATAGGAAAGGAGTACGTCAAGGCTGT 69054
Qy 1898 ATATGTACACCGGCTGTTTAACTTCTATGCAGAG-ACATCATGAGAAACGCTGGGCTG 1956
Db 69055 ATATGTACACCGCTGTTTAACTTCTATGCAGAGTACATCATGAGAAACACTGGGCTG 69114

Qy 1957 GAAGAGCACAAGCTGGATCAAGATTCGCGGAGAAATAGCAATAACCTCAG-ATATGC 2015
Db 69115 GAAGAGCACAAGCTGGATCAAGATTCGCGGAGAAATATCAATAACCTCAGAATATGC 69174
Qy 2016 AGATGATACCACTTATGGCAGAAAGTGAAGAGAACTAAAGAGCTCTTCTGATGAAGT 2075
Db 69175 AGATGACACACCTTATGGCAGAAAGTGAAGAGAACTAAAGAGCTCTTCTGATGAAGT 69234
Qy 2076 GAAGAGGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTTCAAGAAACGAAGATCATGC 2135
Db 69235 GAAGAGGAGAGTGAAGAAATTTGGCTTAAAGCTCAACATTTCAAGAAACGAAGATCATGC 69294
Qy 2136 ATCTGCTCCATCACTTATGGGAAATAGATGGGAAACAGTGGAAACAGTGTCAAGCTT 2195
Db 69295 ATCTGCTCCATCACTTATGGGAAATAGATGGGAAACAGTGGAAACAGTGTCAAGCTT 69354
Qy 2196 TATTTTGGGGGCTCCAAAATCACTGCAGATGGTCACTGCAGCATGAAATTTAAAAGA 2255
Db 69355 TATTTTGGGGGCTCCAAAATCACTGCAGATGGTCACTGCAGCATGAAATTTAAAAGA 69412
Qy 2256 CACTTACTCTTGGAGAAAAAGTTA-ACCAACCTAGATAGCATATTTGAAAGCAGAGAC 2313
Db 69413 CGTTCTCTCTTGGAGGAAAGTTATGACCAACCTAGATAGCATATTTGAAAGCAGAGAC 69472
Qy 2314 ATTACCTTCCACAAAGCCCATCTAGTCAAGGCTATGCTTTTCCAGTGTCAATGAT 2373
Db 69473 ATTACCTTCCACAAAGGCTCATCTAGTCAAGGCTATGCTTTTCCAGTGTCAATGAT 69532
Qy 2374 GGATGTGAGAGTTGGACTGTGAAGAAAGCTGAGCACTGAAGAAATGATCTTTTGAAGT 2433
Db 69533 GGATGTGAGAGTTGGACTGTGAAGAAAGCTGAGCACTGAAGAAATGATCTTTTGAAGT 69592
Qy 2434 TGGTGTGAGAGAC----- 2449
Db 69593 TGGTGTGAGAGACNN 69652
Qy 2450 -----TCTTGAGAGTCCCTTGGACTGCAAGGAGATCCAA 2483
Db 69653 NNN 69712
Qy 2484 CCAAGTCCATCTGAAGAGATCAGCCCTGGATTTCTTTGGAAGGAATGATGCTAAAGCT 2543
Db 69713 CCAAGTCCATCTGAAGAGATCAGCCCTGGATTTCTTTGGAAGGAATGATGCTAAAGCT 69772
Qy 2544 GAACTCCAGTACTTTGGCCACCTGATCAGAGAGCTGACTCACTGGAAGAACCCCTGAT 2603
Db 69773 GAACTCCAGTACTTTGGCCACCTGATGAGAGAGTTGACTCTTGGAAAGAA-CTGAT 69830
Qy 2604 GCTGGAGGAGTTGGGGCAGGAGAGAGGGGACGACAGAGATGAGATGGCTGATGG 2663
Db 69831 GCTGGAGGAGTTGGGGCAGGAGAGAGGGGACGACAGAGATGAGATGATGGATGG 69890
Qy 2664 CATCACTGACTGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2723
Db 69891 CATCACTGACTGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 69950
Qy 2724 GGCCTGTCTGCGCGCATTCATGGGCTCAAGAGTGGACACGACTGAGCAACTGAAC 2783
Db 69951 AGCTGGGCTGCTGCGATTCATGGGCTCGAAGAGTGGACACGACTGAGCAACTGAAC 70010
Qy 2784 TGAATGAACTGTACTGAA 2802
Db 70011 TGAATGAACTGAA 70029

RESULT 27

AC107065
LOCUS AC107065 175949 bp DNA linear NAM 18-SEP-2003
DEFINITION Bos taurus clone rp42-513g13, complete sequence.
ACCESSION AC107065
VERSION AC107065.5 GI:34849950
KEYWORDS HTG.

SOURCE	Bos taurus (cow)	Db	86668	GTATTCTGCTTGGAGACCCCAATGAAACAGTATGAAAAGCGCAAAATGAAAGGATCTACTGAA	86727
ORGANISM	Bos taurus	Qy	393	AGAGGAACCTCCCGAGTCTAGTAGTGGCCCAATATGCTACTGGAGATCAGTGGAGAAATAA	452
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea; Bovidae; Bovinae; Bos.	Db	86728	AGAGGAACCTCCCGAGTCTAGTAGTGGCCCAATATGCTACTGGAGATCAGTGGAGAAATAA	86787
AUTHORS	1 (bases 1 to 175949)	Qy	453	CTCCAGAAAGATGAAGAGATGGAGCCCAAGCAAAAGAAATACCAGCTGTGGATGTGAC	512
TITLE	Bos Taurus BAC Clone rp42-513g13	Db	86788	CTCCAGAAAG-----GATGGAGCCAAAGCAAAACCAATCCAGTGTGTGGATGTGAC	86899
JOURNAL	Unpublished	Qy	513	TGCTGATATAAGCAAGTCCGATGCTGTAAAGAGCAATATTGCTATAGGAACCTCGGAATGT	572
AUTHORS	Hudson,H., Yoon,Y., Shaull,S., Yao,Z., Seroussi,E. and Roe,B.A.	Db	86840	TGCTGATAGAGCAAGATCTGATGCTATAAAGAGCAGTATTGCTATAGGAACCTCGGAATGT	86899
TITLE	Direct Submission	Qy	573	CAGTCCATGAATCAAGGCAAAATTTGAAGTGGTCAAAAGAGATGGCAAGAGTGAATGT	632
JOURNAL	Submitted (15-JAN-2002) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Db	86900	TAGGTCCATGAATCAAGGCAAAATTTGAAGTGGTCAAAAGAGAGCGCAAGAGTGAACAT	86959
REFERENCE	3 (bases 1 to 175949)	Qy	633	CAACATTTCTAGGAATCAGCGAACTAAATGGACTGGAATGGGTGAATTTAACTCAGATGA	692
AUTHORS	Hudson,H., Yoon,Y., Shaull,S., Yao,Z., Seroussi,E. and Roe,B.A.	Db	86960	CGACGTTCTAGGAATCAGTGAATGAATGGACTGGAATGGGTGAATTTAACTCAGATGA	87019
TITLE	Direct Submission	Qy	693	CCATTATATCTACTACTCGGGCAAGAAATCCCTCAGAGAAATGAGTAGCCATCATGGT	752
JOURNAL	Submitted (17-SEP-2002) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Db	87020	CCATTATATCTACTACTCGGGCAAGAAATCCCTCAGAGAAATGAGTAGCCATCATGGT	87079
REFERENCE	5 (bases 1 to 175949)	Qy	753	CAACAAAGAGTCTGAAATGGCAATCTGGATGGCAATCTCAAAAACGACAGAACCATCTC	812
AUTHORS	Hudson,H., Yoon,Y., Shaull,S., Yao,Z., Seroussi,E. and Roe,B.A.	Db	87080	CAACAAAGAGTCTGAAATGGCAATCTGGATGGCAATCTCAAAAACGACAGAACCATCTC	87139
TITLE	Direct Submission	Qy	813	TGTTGTTTCCAAAGCAAAACCAATTCATATACAGTAATCCAAAGTCTATGCCCAACCCAG	872
JOURNAL	Submitted (18-SEP-2003) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Db	87140	TGTTGTTTCCAAAGCAAAACCAATTCATATACAGTAATCCAAAGTCTATGCCCAACCCAG	87199
COMMENT	On Sep 18, 2003 this sequence version replaced gi:22954011. Center: Department Of Chemistry And Biochemistry The University Of Oklahoma Center code:UOKNOR	Qy	873	TAATGCTGAGAGAGCTGAAGTTGAACGGTCCCTATGAAGACCTACAGAACCTTTTAGAAT	932
FEATURES	Location/Qualifiers	Db	87200	TAACGCTGAAGAGCTGAAGTTGAACGGTCCCTATGAAGACCTACAGAACCTTTTAGAAT	87259
source	1. 175949	Qy	933	AAACACCCAAAAAGATGTCCTTCTCATTTATAGGGAGCTGGAAATGCAAAAGTAGGAAGCA	992
	/organism="Bos taurus"	Db	87260	AAACACCCAAAAAGATGTCCTTCTCATTTACAGGGAGCTGGAAATGCAAAAGTAGGAAGCA	87319
	/mol_type="genomic DNA"	Qy	993	AGAAACACCTGGAGTAACAGGCAAAATTTGGCCCTTGGAAATACGGAATGAAGAGGCAAG	1052
	/db_xref="taxon:9913"	Db	87320	AGAAAGACCTGGAGTAACAGGCAAAATTTGGCCCTTGGAAATACGGAATGAAGAGGCAAG	87379
	/clone="rp42-513g13"	Qy	1053	ACTAATAGAGTTTTCGCAAGAAATGCACTGTGTATAGCAAAACCCCTCTTCCCAACAACA	1112
	/clone_lib="RPCI - 42 Male Bovine BAC Library"	Db	87380	GCTAATAGAGTTTGGCAAGAGATGCACTGTGTATAGCAAAACCCCTCTTCCCAACAACA	87439
Query Match	56.6%; Score 2022.6; DB 1; Length 175949;	Qy	1113	CAAGAGAAAGCTCTACACATGACATCACAGATGGTCAACACCGAAATCAGATTGATTA	1172
Best Local Similarity	75.6%; Pred. No. 0;	Db	87440	CAAGAGAAAGCTCTACACATGACATCACAGATGGTCAACACCTGAAATCAGATTGATTA	87499
Matches 2541; Conservative	0; Mismatches 194; Indels 624; Gaps 9;	Qy	1173	TATTCCTTTTCAGCCCAAGAGTGGAGAGCTCTATACAGTCAAGCAAAACAGACGAGGAC	1232
Qy	35 CCTCGTCAAGGTAAGGAGCAGTAGCTGCGCTTTGCTGGAGCAGCCGTAAGAGATACCC	Db	87500	TATTCCTTTTCAGCCCAAGAGTGGAGAGCTCTATACAGTCAAGCAAAACAGACGAGGAC	87559
Db	86370 CCTCATCAAGGTAAGGAGCAGTAGCTGCGCACTTTGCTGGACCAAGCCGTAAGAGATACCC	Qy	1233	TTACTGTGGCTCAGATCATGAACCTCTTTATTGCCAAATTTAGACTTTAAATTCAGAAAGT	1292
Qy	95 CACGCCCAAGTAAGAGAACCCCAAGTAAGTGTAGTGTGTGAGAGGGCTACAGAG	Db	87560	TGACCGTAGTCTCAGATCATGAACCTCTTTATTGCCAAATTTAGACTTTAAATTCAGAAAGT	87619
Db	86430 CACGTCCAAGTAAGAGAACCCCAAGTAAGCAGTGTGTGAGAGGGCTACAGAG	Qy	1293	AGGGAACCACTAGATCAGTCAAGTAAGAGCTTAATCCATCCCTTTATGATTATACAGT	1352
Qy	155 GCAGACATCTGAAACCATACACGCAAGAACTAGTCAATCTAATCAGTACAGTACCA	Db	87620	AGGGAACCACTAGATCAGTCAAGTAAGAGCTTAATCCATCCCTTTATGATTATACAGT	87679
Db	86490 GCAGACATCTGAAATCATATATACAGAAACTAGCCAACTGATCAGAC--GGACCA	Qy	1353	GGAAAGTGAAGTAAGATTTAAAGGGCTAGATCTGTATAGACAGAGTACCTAATGAACATG	1412
Qy	215 GCCTTGTCTAATCAAGTAAGCAATGCGCC--CGTGGGGCAACCAAGATGGGAGGT	Db	87680	GGAAAGTGAAGTAAGATTTAAAGGGCTAGATCTGTATAGACAGAGTACCTAATGAACATG	87739
Db	86548 GCCTTGTCTAATCAAGTAAGCAATGCGCC--CGTGGGGCAACCAAGATGGGAGGT	Qy	1413	GACAGAGTTCATGACATTTGACAGGAGAGGATCGAGCCATCCCATCGGAAAGAA	1472
Qy	274 CATGGTGGAGAGTCTGACAGAACTGTGCTCAGTGGAGAGGGAAT--GCAAAACCACTTCA	Db	87740	GACTGAGGTTTCATGACATTTGACAGGAGAGGATCAAGACCATCCCATCGGAAAGAA	87799
Db	86608 CATGGTGGAGAGTCTGACAGAACTGTGCTCAGTGGAGAGGGAATGGCAAAACCACTTCA	Qy			
Qy	333 GTATTCTTGGCTTGAAGACCCCATGAAACAGTATGAAAGGCAAAATGATAGGATCTGAA	Db			

QY	1473	ATGCAAAAAGCAAAATGCGTCTCTGGGAGGCCCTTACAAATAGCTGTGAAAAGAGAGA	1532
Db	87800	ATGCAGAAAGCAAAATGCGTCTCTGGGAGGCCCTTACAAATAGCTGTGAAAAGAGAGA	87859
QY	1533	AGTGAAAAGCAAAAGGAAAAGAGATATAAAGCATCTGAATCAGAGTTCCAAA---	1588
Db	87860	AGCAAAAAGCAAAAGGAAAAGAGATATAAAGCATCTGAATCAGAGTTCCAAAAGAC	87919
QY	1589	-----	1588
Db	87920	AGCAAGAGAGATAAGAAAGCCTTCTCTCAGCATCAATGCAAGAAATAGAGGAAAAACA	87979
QY	1589	-----	1588
Db	87980	CAGAATGGGAAACACTTGAGATCTCTTCAAGAAATATTAGATACCAAAGAAACATTTCA	88039
QY	1589	-----	1588
Db	88040	TGCAAGACGGGCTTGATAAAGCAGAGAAATCATATGGACCTAACAGAAAACAGAGATAT	88099
QY	1589	-----	1588
Db	88100	TAAGAAGAGTGCAAGAATAACAGAGAACTGTACAAAAAGATCTTCAGACCAAGA	88159
QY	1589	-----	1588
Db	88160	TAATCATGTGTGTGATCGCTGACCTGGAGCCAGACATCTCTGGAATGTGAAGTCAAGTG	88219
QY	1589	-----	1588
Db	88220	GGCCTTAGAAGCATCACTACGAACAAAGCTAGTGGAGGTGATGGAATTCAGTTACGCT	88279
QY	1589	-----	1588
Db	88280	ATTTACATCTGAAAGATGATGCTGTGTGAAGTGTGCACTCAATTTGCCAGCAAACTCG	88339
QY	1589	-----	1588
Db	88340	GAAAACTCAGCGGTGCCACAGGCTGGAAAAGTCACTTTTCATTCOAATCCCAAGAA	88399
QY	1589	-----	1588
Db	88400	AGGCAATGCCAAGAATGCTCAAACTACTGCACAATTTGTACTCATCTCACAGCTAGTAA	88459
QY	1589	-----GAACTTCAGT	1599
Db	88460	AGTAATGCTCAAAATTTCTCAGGCCAGGCTTACGAAAAAGTGAATCTGTGAACCTCCAGA	88519
QY	1600	TGTTCAAGCTGGTTTTAGAAAAGTACAGAGAACAGAGACCAAAATGCGCAACCTCCTCTG	1659
Db	88520	TATTCAGCTGGTTTTAGAAAAGCGAGAGAACAGAGATCAAAATGGCAACATCCACTG	88579
QY	1660	TATCATGGAAAAAGCAAGAGAGTCCAGAAAAACATCTATTCTCTGCTTTATGACTATGC	1719
Db	88580	GATCATCGAAAAAGCAAGAGAGTTCAGAAAAACATCTATTCTCTGCTTTTGTGACTATGC	88639
QY	1720	AAAAGCCTTTGACTGTGGGGTCAACAATAAATGTCGAAAAATCTCGAAAAGGATGGGAAT	1779
Db	88640	CAAAGCTTTTGAATGTGTGGATCACAGTAACTGTGGAANAATCTGAAGAAGATGGGAAT	88699
QY	1780	ACCAGACCACTGACTGACTCTTTGAAAAATTTGTTATGCAAGTCAAGGAAGCAAGTTAG	1839
Db	88700	ACCAGACCACTGATCTGCTCTTTGAAAAAGCTTATATGCAAGTCAAGGAAGCAACAGTTAG	88759
QY	1840	AAC-TGGACATGGAAACAACAGACTGTTCCAGTACGAAAAAGAGATGTCTCAAGGCTGTA	1898
Db	88760	AACTTGACATGGAAACAACAGACTGTTCCAAATAGAAAAAGAGTACGTGAGGCTGTA	88819
QY	1899	TATTGTCAACCGGCTGTTTTAACTTCTATGACAG-ACATCATGAAAAACGCTGGGCTGG	1957
Db	88820	TATTGTCACTCTGCTTTATTTAACTGATATGCAAGTACATCATGAAAAACGCTGGGCTGG	88879

QY	1958	AAGAGCACAAGCTGGAATCAAGATTGCCGGGAGAAATAGCAATAACTCCAGATATGCAG	2017
DB	88980	AAGAAGCACAAGCTGGAATCAAGATTGCCAGGAGAAATATCAATAACTCCAGATATGCAG	88939
QY	2018	ATGATACCCACCTTTATGGCAGAAAGTGAAGAGAACTAAAAAGCCTCTTTGATGAAGTGCA	2077
DB	88940	ATGACACCACTTTATGGCAGAAAGTGAAGAGAACTAAAAAGCCTCTTTGATGAAGTGCA	88999
QY	2078	AAGAGGAGAGTGAANAAGTTGGCTTAAAGCTCAAATTACAGAAAAACGAAGATCATGGCAT	2137
DB	89000	AAGTGGAGAGTGAANAAGTTGGCTTATAGCTCAAATTACAGAAAAACGAAGATCATGGCAT	89059
QY	2138	CTGGTCCCACACTTCATGGGAAAATAGATGGGAAAACAGTGAACACAGTGCACAGACTTTTA	2197
DB	89060	CTGGTCCCACACTTCATGGGAAAATAGATGGGAAAACAGTGAACACAGTGCAGACTTTTA	89119
QY	2198	TTTTTTTGGGGGGCTCCAAAATCAGTCAGATGGTGAAGCTGCAGCCCATGAAATTTAAAGACA	2257
DB	89120	TTTTTTTGGGGCTCCAAAATCAGTCAGATGGTGAAGCTGCAGCCCATGAAATTTAAAGCGC	89164
QY	2258	CTTACTCCTTTGGAGAAAAAGTTA--ACCAACCTAGATAGCATATTGAAAAAGCAGAGACAT	2315
DB	89165	CTTACTCCTTTGGAGAGGGGTATGACCAACCTAGATAGCATATTGAAAAATCAGAGACAT	89224
QY	2316	TACCTTGCCAAACAAAGCCCATCTAGTCAAGGCTATGGTTTTTCAGATGGCTCATGTATGG	2375
DB	89225	TACTTTGCCAAACAAAGTCTGTCTAGTCAAGGCTATGGTTTTTCAGAGGCTCATGTATGG	89284
QY	2376	ATGTGAGAGTTGGACTGTGAAGAAAGCTGAGCAGCTGAAGAAATGTATGTTTTGAACTGTG	2435
DB	89285	ATGTGAGAGTTGGACTGTGAAGAAAGCTGAGTCCGGAAGAAATTGATGTTTTGAACTGTG	89344
QY	2436	GTGTTGGAGAAAGACTCTTGAGAGTCCCTTTGGACTGCAAGAGAGATCCAACAGTCCATCT	2495
DB	89345	GTGTTGGAGAAAGACTCTTGAGAGTCCCTTTGGACTGCAAGAGAGATCCAACAGTCCATCT	89404
QY	2496	GAAGGAGATCAGCCCTGGGATTTCTTTTGAAGGAATGATGCTAAAGCTGAAACTCCAGTA	2555
DB	89405	AAAGGAGATCAGTCTCTGGGTGTTCTTTTGAAGGAATGATGCTGAACTGAAACTCCAGTA	89464
QY	2556	CTTTGGCCACCTGATCAGAAAGACTGACTCACTGGAAAAAGACCTTGATGCTGGAGGGAT	2615
DB	89465	CTTTGGCCACCTCATGCGAAGAGTTGACTCAITGGAAAAAGAGCGCTGATGCTGGAGGAT	89524
QY	2616	TGGGGGAGAGGAGAGGGGAGCAAGAGGATGAGATGGCTGGATGGCATCACCTGACTC	2675
DB	89525	TGGGGGAGAGGAGAGGGGAGCAACTGAGGATGAGATGGTGGATGGCATCACCAGCTG	89584
QY	2676	GATGCAAGTGAAGTCTGGGTCAACTCCTCGAGTTGGTGATGGACAGGGAGCGCTGCTGC	2735
DB	89585	GATGCAAGTGAAGTCTGAGTGAACTCTTGGAAGTTGGTGATGGACAGGGAGCGCTCTCGTGA	89644
QY	2736	GGCGATTATGGGGTCAAAAAGAGTTGGACACGAGCTGAGCAACTGAACTGAAGTGAAGCT	2794
DB	89645	TGCAATTATGGGATGCCAAAGAGTGCAGACAGCAAGCTGAGCAACTGAACTGAAGTGAAGCT	89703

RESULT 28

AC109914
LOCUS

DEFINITION

ACCESSION

VERSION
KEYWORDS

SOURCE
ORGANIS

STAFF

REFERENCE

AUTHORS

Cariaga, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Lalic, P., Lee-Lin, S.-Q., Legaspi, R., Maduro, O.L., Maduro, V.B., Margulies, E.H., Mastello, C., Maskeri, B., McDowell, J., Paguirigan, C., Pearson, R., Portnoy, M.E., Prasad, A., Reddix-Dugue, N., Schander, K., Schueler, M.G., Sison, C., Stantryp, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L., Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 177535)
Green, E.D.

TITLE
JOURNALREFERENCE
AUTHORS

Direct Submission

TITLE
JOURNALREFERENCE
AUTHORS

Direct Submission

TITLE
JOURNAL

COMMENT

Submitted (08-FEB-2002) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 177535)
Green, E.D.

Submitted (04-DEC-2002) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
On Dec 4, 2002 this sequence version replaced gi:18640643.

----- Genome Center
Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: <http://www.nisc.nih.gov>

Contact: nisc_zooenhgri.nih.gov

----- Project Information

Center project name: cbc

Center clone name: 262G19

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 176756 bases at least Q40
Consensus quality: 177014 bases at least Q30
Consensus quality: 177134 bases at least Q20
Insert size: 146000; agarose-fp
Insert size: 177235; sum-of-contigs
Quality coverage: 12.93x in Q20 bases; agarose-fp
Quality coverage: 10.65x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

* 1 25820: contig of 25820 bp in length
* 25821 25920: gap of unknown length
* 25921 69500: contig of 44030 bp in length
* 69501 70050: gap of unknown length
* 70051 170904: contig of 100854 bp in length
* 170905 171004: gap of unknown length
* 171005 177535: contig of 6531 bp in length.

FEATURES
source

1. 177535

/organism="Bos taurus"

```

/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="RP42-262G19"
/clone_lib="RP42"
1. 25820
/note="assembly_fragment
clone_end:SP6
vector_side:left"
1. 18394
/note="clone overlaps with GenBank Accession Number
AC108991 clone RP42-252K20 (center project name cbb)"
25921...69950
/note="assembly_fragment"
70051...170904
/note="assembly_fragment"
150522...177535
/note="clone overlaps with GenBank Accession Number
AC108913 clone RP42-156J16 (center project name cbl)"
171005...177535
/note="assembly_fragment
clone_end:T7
vector_side:right"

Query Match      56.5%; Score 2018.9; DB 1; Length 177535;
Best Local Similarity 75.7%; Pred. No. 0;
Matches 2572; Conservative 0; Mismatches 201; Indels 625; Gaps 12;

QY      2 TCAGGAAGGGCGGCGAGTGAGGAGGTACTACCCCTCGTCCAAAGGTAAGGAGCAGTAGCT 61
Db      122254 TCAGGAGGGCGGCGGTGAGGAGATAT-----CCCTGTGTCAGGAGCAGTGCT 122308

QY      62 GCGCTTGTGCGAGCAGCGGTAAAGAGATACCCGAGCGCCAGGTAAAGAAACCCAGT 121
Db      122309 GTGCTTTCAGGAGCAGCTGTGAAGAGATACCCACATCCAAAGTAAGAGAAACCCAAAT 122368

QY      122 AAAGTGGTAGGTGTTGTGAGAGGGCATCAGAGGGCAGACATACCTGAAACCATACACGAG 181
Db      122369 AAGTGGTAGGTGTTGCAAGAGGGCATCAGAGGGCAACACACTGAACCATATTTCAG 122428

QY      182 AAACTAGTCAATCTAATCACACTAGGACACAGGCTTGTCTAATCTAATGAATAGC 241
Db      122429 AAACTAGTCAATCTAATCACACTAGGACCATAGCTTGTCTAATCTAATGAATAGC 122488

QY      242 CATGCCGTGGGGCAACCAAGATGGCGAGGTCTGTTGGAGAGATCTGACAGAAATGTGG 301
Db      122489 CATGCTGTGGGGCCACCACAGACGGGGCGGTGTCATGTTGGAGAGATCTGACAGAAATGTGG 122548

QY      302 TCACCTGGAGAGGGAAT-GCAAAACCATTCAGTATTCTTGCTTTGAGAACCCCATGAAC 360
Db      122549 TCACCTGGAGAGGGAATAGCAACCACTTCAGTATTCTTGCTTTGAGAACCCCATGAAC 122608

QY      361 AGTATGAAAGGCAAAATGATAGGATCTGAAAGAGGAACCTCCCGAGTCTAGTAGTGCC 420
Db      122609 AGTATGAAAGGCAAAATGATAGGATCTGAAAGAGGAACCTCCCGAGTCTAGTAGTGCC 122668

QY      421 CCATATGCTTCTGGAGATCAGTGGAGAAATAACTCCAGAAAGAAATGAAGAGATGGAGCCA 480
Db      122669 CAATATGCTTCTGGAGATCAGTGGAGAAATAACTCCAGAAAGAAATGAAGAGATGGAGCCA 122728

QY      481 AGCAAAAAGATACCCAGCTGTGGATGTGACTGTGTGATATAGCAAGGTCCGATGCTGT 540
Db      122729 AGCAAAAAGCAATACCTAGTGTGGATGTGACTGTGTGATATAGCAAGGTCCGATGCTGT 122788

QY      541 AAAGAGCAATATTGTCATAGGAACCTGGAAATGTGAGGTCCATCAATCAAGGCAAAATGGAA 600
Db      122789 AAAGAGCAATATTGTCATAGGAACCTGGAAATGTGAGGTCCATCAATCAAGGCAAAATGGAA 122848

QY      601 GTGTGCAAAAGAGATGCGAAGAGTGAATGTCAACATTCTAGGAATCAGCGAATATAAA 660
Db      122849 GTGTGCAAAAGAGATGCGAAGAGTGAATGTCAACATTCTAGGAATCAGCGAATATAAA 122908

QY      661 TGGACTGGAATGGGTGAATTTAACTCAGATCACCATTATATCTACTCTGCGGCGCAGGAA 720

```

Db	122909	TGGACTGGAATGGGTGAATTTAACTCAGATGACCAATTATATCTACTACTCGGGCAGGAA	122968
Qy	721	TCCCTCAGAGAAATGGAGTAGCCCATCATGTGTCAACAAAAGAGTCCGAAATGCGAGTACTT	780
Db	122969	TCCCTCAGAGAAATGGAGTAGCCCATCATGTGTCAACAAAAGAGTCCGAAATGCGAGATTT	123028
Qy	781	GGATGCACTCTCAAAAACGACAGAGATGATCTCTGTTTGGTTTCCAAAGGCAAAACCAATTCAT	840
Db	123029	GGATGCAATCTCAAAAACGACAGAGATGATCTCTGTTTCCAAAGGCAAAACCAATTCAT	123088
Qy	841	ATCAGATTAATCCCAAGTCTATGCCCCCAACCACTAGTAACTGCTGAAGAGCTGAAGTTGAACGG	900
Db	123089	ATCAGATTAATCCCAAGTCTATGCCCCCAACCACTAGTAACTGCTGAAGAGCTGAAGTTGAACGG	123148
Qy	901	TCCTATGAAGACCTTACAAGACCTTTTATGAACCTAACACCCCAAAAAGATGCTCTTCTCAT	960
Db	123149	TTCTATGAAGACCTTACAAGACCTTTTATGAACCTAACACCCCAAAAAGATGCTCTTCTCAT	123208
Qy	961	ATAGGGGACTGGAATGCAAAATAGTAGAGCAAGCAACACCTGAGTAAACAGGCAAAATTT	1020
Db	123209	ATAGGGGACTGGAATGCAAAATAGTAGAGCAAGCAACACCTGAGTAAACAGGCAAAATTT	123268
Qy	1021	GGCCTTGGAAATACGGAATGAAGCAGGGCAAAAGACTAATAGAGTTTTCGCAAGAAATGCA	1080
Db	123269	GGCCTTGGAAATACGGAATGAAGCAGGGCAAAAGACTAATAGAGTTTTCGCAAGAAATGCA	123328
Qy	1081	CTGCTCATAGCAAAACACCTCTTCCAAACACACAGAGAGACTCTACACATGACATCA	1140
Db	123329	CTGCTCATAGCAAAACACCTCTTCCAAACACACAGAGAGACTCTACACATGACATCA	123388
Qy	1141	CCAGATGCTCAACACCCGAATCAGATTGATTATATCTTTCAGCCCAAAAGATGAGAAAGC	1200
Db	123389	CCAGATGCTCAACACCCGAATCAGATTGATTATATCTTTCAGCCCAAAAGATGAGAAAGT	123448
Qy	1201	TCTATACAGTCAGCAAAACCAAGACCCAGGACTTACTGTGGCTCAGATCATGAACTCCTT	1260
Db	123449	TCTATACAGTCAGCAAAACCAAGACCCAGGACTTACTGTGGCTCAGATCATGAACTCCTT	123508
Qy	1261	ATTGCCAAATTCAGACTTAAATTCGAAGAGTAGGGAAACCACTAGATCACTCAGGTAA	1320
Db	123509	ATTGCCAAATTCAGACTTAAATTCGAAGAGTAGGGAAACCACTAGATCACTCAGGTAA	123568
Qy	1321	GACCTAAATCCCAATCCCTTATGATTATACAGTGGAAAGTGAGAAATAGATTAAAGGCTTA	1380
Db	123569	GACCTAAATCCCAATCCCTTATGATTATACAGTGGAAAGTGAGAAATAGATTAAAGGCTTA	123628
Qy	1381	GATCTGATAGACAGATACCTTAATGAACACTATGACAGAGGTTCTATGATGTACAGGAG	1440
Db	123629	GATCTGATAGACAGATACCTTAATGAACACTATGACAGAGGTTCTATGATGTACAGGAG	123688
Qy	1441	ACAGGGATCGAGACCATCCCATGGAAAAGAAATGTC-AAAAAAGCAAAATGGCTGTCTGG	1499
Db	123689	ACAGGGATCGAGACCATCCCGTGGAAAAGAAATGCAAAAAGCAAAATGGCTGTCTGG	123748
Qy	1500	GGAGGCTTACAATAGCTGTGAAAAGAGAGAGTGAAGAAGCAAAAGGAAAAG	1559
Db	123749	GGAGGCTTACAATAGCTGTGAAAAGAGAGAGTGAAGAAGCAAAAGGAAAAG	123808
Qy	1560	ATAAAGACATCTGAATGCAGAGTTCCAAA-----	1588
Db	123809	ATAAAGACATCTGAATGCAGAGTTCCAAAAGATAGCAAGAGATGAAGAAGCCCTCTT	123868
Qy	1589	-----	1588
Db	123869	CAGCGATCAATGCAAAAGAAATAGAGGGAAGCAACAGAAATGGGAAAGACTAGAGATATCTT	123928
Qy	1589	-----	1588
Db	123929	CAAGAAATTAGAGATACCAAGGGAACATTAAACGTAAGATGGGCTCGATAAAGGCGAG	123988
Qy	1589	-----	1588
Db	123989	AAATGGTATGGACCTTAACAGAGCAGAGAGATTTAAGAGAGATGGCAAGATACACAGA	124048
Qy	1589	-----	1588
Db	124049	AGAACTGTACAAAAAGTTCTTCAAGACCAAGATAATCACAATGGTGTGATCACTGACCT	124108
Qy	1589	-----	1588
Db	124109	AGAGCCAGACATCTCTGGAATGTGAAGTCAAGTGGGCTTTAGAAAGCATCACTACGAACAA	124168
Qy	1589	-----	1588
Db	124169	AGCTAGTGGAGGTGATGGAATTCAGATTCCAGCTCTTTCAAACTCCTGAAAGATGATGCTGT	124228
Qy	1589	-----	1588
Db	124229	GAAAGTGTACACTCAATATGCGAGCAAAATTTGCAAACTCAGCAGTGGCCACAGGACTG	124288
Qy	1589	-----	1588
Db	124289	GAAAACGTCAGTTTTCATTCCTCAATCCCAAGAAAGGCAATGCCAAAGAAATGCTCAAACTA	124348
Qy	1589	-----	1588
Db	124349	CTGCACAAATTTGACTCATCTCAGATGATAGTAAAGTAAATGCTCAAAATTTCTCCAGCCAG	124408
Qy	1589	-----	1588
Db	124409	GCTTCAGCAATATGTGAACCGTGAACTTCCAGATGTTCAAGCTGGTTTTAGAAAAAGCGAG	124468
Qy	1589	-----	1588
Db	124469	AGAAACGACAGACCAAAATTTGCCAACATCTCTGTATCATGGAAGAAAGCAAGAGAGTTCA	1686
Qy	1627	AGAAACGACAGACCAAAATTTGCCAACATCTCTGTATCATGGAAGAAAGCAAGAGAGTTCA	1686
Db	124469	AGAAACGACAGACCAAAATTTGCCAACATCTCTGTATCATGGAAGAAAGCAAGAGAGTTCA	124528
Qy	1687	GAAAAACATCTATTTCTGCTTTTATTGACTATGCAAAAGCCTTTGACTGTGGGGGTCACAA	1746
Db	124529	GAAAAACATCTATTTCTGCTTTTATTGACTATGCAAAAGCCTTTGACTGTGGGGGTCACAA	124584
Qy	1747	TAACTGTGAAAAATTTCTCAAGAGGATGGAATACCAAGACCACTGACCTGACTCTTTGAA	1806
Db	124585	TAACTGTGAAAAATTTCTCAAGAGATGGAATACCAAGACCACTGACTGATGCTCTTTGAG	124644
Qy	1807	AAATTTGATGAGGTCAGAGTCAAGAACCACTGTAGAACTGGCAATGGAACCAAGAGCTGGT	1866
Db	124645	AAACCTGTATGAGGTCAGAGTCAAGAACCACTGTAGAACTGGCAATGGAACCAAGAGCTGGT	124704
Qy	1867	CCAACTAGGAAAAGAGTATGTCAAGGCTGTATATGTCAACCCGCTTGTAACTCTTA	1926
Db	124705	CCAACTAGGAAAAGAGTCCATCAAGGCTGTATATGTCAACCCGCTTGTAACTCTTA	124764
Qy	1927	TGCAGAG-ACATCATGAGAAACGCTGGGCTGGAAGAACCAAGCTGGAATCAAGATTGC	1985
Db	124765	TGCAGAGCAGCTCATGAGAAATGCTGGGCTGGAAGAACCAAGCTGGAATCAAGATTTC	124824
Qy	1986	CGGAGAAAATAGCAATAACCTCAGATATGCAAG-TGATACCACTCTTATGGCAGAAAGTG	2044
Db	124825	TGGGAGAAAATATCAATTAACCTCAGATATGCAAGTAAACCACTCTTATGGCAGAAAGTG	124884
Qy	2045	AGAGAGAACTAAAAAGCCTTTGATGAGGTCGAAAGAGAGAGTGAAGAGTTGGCTTAA	2104
Db	124885	AGAGAGAACTAAAAAGCCTTTGATGAGGTCGAAAGAGAGTGAAGAGTTGGCTTAA	124944
Qy	2105	AGCTCAACATTCAGAAAACGAAAGATCATGGCATCTGGTCCCATCTTTCATGGGAAATAG	2164
Db	124945	AGCTCAACATTCAGAAAACGAAAGATCATGGCATCTGGTCCCATCTTTCATGGGAAATAG	125004
Qy	2165	ATGGGAAAACAGTGGAAACAGTGTGAGCTTTTATTTTGGGGGCTCCAAATCACTGC	2224
Db	125005	ATGGGAAAACAGTGGAAATAGTGTGAGGTTTATTTTGGGGGCTCCAAATCACTGC	125057
Qy	2225	AGATGCTGACTGCAGCCATGAAATTAAGAGACACTTACTCTTTGGAAGAAAAGTTA--AC	2282
Db	125058	-----ACTGCAGTCATGAAATTAAGAGATGC-TACTCTTTTGGAGAAAAGTTATGAC	125108

Db	166527	CCAGTCTAGTAGGGCCCCAATATGCTACTGAGATCACTGGAGAAATAACTCCAGAAAGA	166586	Db	167605	GCAAAATGGCTGTATGGGGAGGCTTTACAAATAGCTGTGAAAAAGAGAGAGTGAAGC	167664
Qy	464	ATGAAGAGATGGAGCCAAAGCAAAAGAAATACCCAGCTGTGATGCTGCTGATATAA	523	Qy	1543	AAAGGAAAAAGGAAAGATTAAGCATCTGAATGCAGAGTTCAAA-	1588
Db	166587	ATGAAGGATGGAACCAAGCAAAACAAATACCCAGCTGTGATGCTGCTGATAGAA	166646	Db	167665	AAAGGGCAAAAGGAAAGATGTAAGCATCTGAATGCAGAGTTCAAAAGAAATAGCCAGAAGA	167724
Qy	524	GCAAGTCCGATGCTGTAAAGAGCAATATTGCTAGGAACCTGGAATGTCAAGTCCCATGA	583	Qy	1589	-----	1588
Db	166647	GCAAGTCCCATGCTGTAAAGGCAATGTTTCATAGAAACCTGGAAATGCAGGCCCATGA	166706	Db	167725	GATAAGAAAGCCTTTCTCATCGATCAGTGCAAAAGAAATAGAGGTAAACAACCTCAATGGGA	167784
Qy	584	ATCAAGGCAAAATGGAAGTGTCAAAAGAGATGGCAAGAGTGAATGTCAACATTTCTAG	643	Qy	1589	-----	1588
Db	166707	ATCAAGGCAAAATGGAAGTGTCAAAAGAGATGGCAAGAGTGAATGTCAACATTTCTAG	166766	Db	167785	AAGACTAGAGATCTCTTCAAGAAATCAGAGATACCAGGGGACATTTTCATGCAAGATG	167844
Qy	644	GAATCAGCGAACTAAATGGACTGGAATGGGTGAATTTAACTCAGATGACCATATATCT	703	Qy	1589	-----	1588
Db	166767	GAATCAGCGAACTGAAATGGACTGGAATGGGTGAATTTAACTCAGATGACCATATATCT	166826	Db	167845	GGCTCGATTAAGGACAGAAATGGTATGGATCTAAACAGAAGCAGAGATATTAAAGGAGGTG	167904
Qy	704	ACTACTGGGCGAGGAATCCCTCAGAGAAATGGAGTAGCCATCATGTCACAAAAGAG	763	Qy	1589	-----	1588
Db	166827	ACTACTGGGCGAGGAATCCCTCAGAGAAATGGAGTAGCCATCATGTCACAAAAGAG	166886	Db	167905	GCAAGAATACACAGAGAACTGTACAAAAGATCTTCACGACCCAGATAATCACGACGG	167964
Qy	764	TCCGAAATGCAGTACTTGGATGCAATCTCAAAAAGCAGAGATGATCTCTTTTCTTCC	823	Qy	1589	-----	1588
Db	166887	TCCGAAATGCAGTACTTGGATGCAATCTCAAAAAGCAGAGATGATCTCTTTTCTTCC	166946	Db	167965	TGTGATCACTGACCTAGGTGAGACATCCTGGAAATGTGAAGTCAAGTGGGCTTAGAAAG	168024
Qy	824	AAGCAAAACCATTCATATACAGTAATCCAAAGTCTATGCCCAACCAAGTAATGCTGAAG	883	Qy	1589	-----	1588
Db	166947	AAGCAAAACCATTCATATACAGTAATCCAAAGTCTATGCCCAACCAAGTAATGCTGAAG	167006	Db	168025	CATCACTACGAAACAAAGCTAGTGGAGGTGATGGAATTCAGTCGAGCTATTTCAAACTCT	168084
Qy	884	AAGCTGAAGTTGAACGCTCTATGAAGACCTACAGACCTTTTGAATTAACACCCAAAA	943	Qy	1589	-----	1588
Db	167007	AAGCTGAAGTTGAATGGTCTTATGAAGACCTACAGACCTTTTGAATTAACACCCAAAA	167066	Db	168085	GAAAGATGATGCTGTAAAAAGTCTGCATCAATATGCCAGCAAAATTTGAAAACTCAGCA	168144
Qy	944	AAGATGCTCTTCTCATATTAGGGACTGGAATGCAAAAGTAGGAAGCAAAAGAACACTG	1003	Qy	1589	-----	1588
Db	167067	AAGATGCTCTTCTCATATTAGGGACTGGAATGCAAAAGTAGGAAGCAAAAGAACACTG	167126	Db	168145	GTGCCCACAGGTTGAAAAAGTCACTTTTCATTCCAATCCCAAGAAAGCAATGCTCA	168204
Qy	1004	GAGTAACAGGCAAAATTTGGCTTGGATACGGAATGAAGAGGGCAAGCAATATAGT	1063	Qy	1589	-----	1588
Db	167127	GAGTAACAGGCAAAATTTGGCTTGGATACGGAATGAAGAGGGCAAGCAATATAGT	167186	Db	168205	AGAATCCTCAAACTACCGCACAAATTCACACTCTCATCTGTAGTAAAGTAATGCTCAA	168264
Qy	1064	TTTGCCAGAAATGCACTGCTATAGCAACACCTCTTCCACACACACAGAGAGAC	1123	Qy	1589	-----	1588
Db	167187	TTTGCCA--AAATGCACTGCTCATAGCAACACCTCTTCCACACACACAGAGAGAC	167244	Db	168265	AATTCCTCAAGCCAGCTTCAGCAATACGTGAACCGTGAACCTTCAGATGTTAAAGCTGG	168324
Qy	1124	TCTACACATGGACATCACCAGATGCTCAACACCGAAATCAGATTATATTTTGGCA	1183	Qy	1612	TTTTAGAAAAGTCAGAGAACCCAGAGACCAAAATTCGCAACATCCTCTGTATCATGAAAA	1671
Db	167245	TCTACACATGGACATCACCAGATGCTCAACACCGAAATCAGATTATATTTTGGCA	167304	Db	168325	TTTTAGAAAAGGCGAGAGAACCCAGATATCAAAATTCGCAACATCCGCTGGATCATGAAAA	168384
Qy	1184	GCCAAAGATGGAAGCTCTATACAGTCAGCAAAACAGACACAGGAGCTTACTGTGGCT	1243	Qy	1672	AGCAAGAGAGTTCCAGAAAAACATCTATTTCTGCTTTATTGACTATGCAAAAGCCTTGA	1731
Db	167305	GCCAAAGATGGAAGCTCTATACAGTCAGCAAAACAGACACAGGAGCTTACTGTGGCT	167364	Db	168385	AGCAAGGAGTTCCAGAAAAACATCTATTTTCTGCTTTATTGACTATACAAAAGCCTTGA	168444
Qy	1244	CAGATCATGAATCTCTTATTCGAAATTCAGACTTTAAATTCGAAGAAAGTGGGAAAAACCA	1303	Qy	1732	CTGTGGGGTCACAATATAACTGTGAAAAATCTGAAAGGATGGGAATACCAGACCACT	1791
Db	167365	CAGATCATGAATCTCTTATTAACCAATTCAGACCTAAATGAAAGAAAGTAGGGAACCA	167424	Db	168445	CTGTGGATCACAATAAAC--TGGACATTTCTGAAAGAGTGGGAATACCAGACCACT	168502
Qy	1304	CTAGATCTCAGGTAGACCTTAATCCAAATCCCTTATGATTAATACAGTGAAGTGAAG	1363	Qy	1792	GACCTGACTCTTGAAAAAATTTGTATGCAAGTCAGAGAGCAACAGTTAGAACTGGACATGG	1851
Db	167425	CTAGACCAATCAGGTATGACCTTAATCCAAATCCCTTATGATTAATACAGTGAAGTGAAG	167484	Db	168503	GATCTGCTCTTTGAGAAATCTGTATGCAAGTCAGAGAGCAACAGTTAGAACTGGACATGG	168562
Qy	1364	ATAGATTTAAGGGCTTAGATCTGTAGAG--CAGAGTACCTTAATGAATGACAGAGGTT	1422	Qy	1852	AACAACAGACTGTTCCCAAGTAGGAAAAAGAGTGTGTCAAGGCTGTATATTGTCAACCGG	1911
Db	167485	TTAGATTTAAGGGCTTAGATCTGTAGAGAGAGAGTGGCTGTATGAATGAGGTT	167544	Db	168563	AACAACAGACTGTTCCCAATAGGAAAAAGAGTGTGTCAAGGCTGTATATTATTCACCTGT	168622
Qy	1423	CATGACATTTGACAGGACAGGATCGAGACCACTCCCATGGAAAAAGAAATGCAAAAA	1482	Qy	1912	CTTGTTTAACTTCTATGACAG--ACATCATGAGAAACGCTGGGCTGGAAAGAACCAAGC	1970
Db	167545	CTGACATTTGACAGGACAGGATCAAGACCACTCCCATGGAAAAAGAAATGCAAAAA	167604	Db	168623	TTTATTTAACTTATATGACAGGTACATCATGAGAAATGCTGACCTGGAAGAAACACAAAC	168682
Qy	1483	GCAAAATGGCTGTCTGGGAGGCCCTTACAAATACCTGTGAAAAAGAGAGTGAAGC	1542	Qy	1971	TGGAATCAAGATTTGCCAGTAGAAAAATATCAATTAACCTCAGATATGACCATCTCT	2030
				Db	168683	TGGAATCAAGATTTGCCAGTAGAAAAATATCAATTAACCTCAGATATGACCATCTCT	168742


```
misc_feature      vector_side:left"
145386..192540
/note="assembly_fragment
clone_end:17
vector_side:right"

Query Match      54.5%; Score 1945.1; DB 1; Length 192540;
Best Local Similarity 74.9%; Pred. No. 0;
Matches 2482; Conservative 0; Mismatches 214; Indels 617; Gaps 10;

QY 89 ATACCCCGCCCAAGGTAAGAGAAACCAAGTAAGATGGTAGGTGGTGGTGGAGGGGCAT 148
Db 17358 ATATCTTATGTCCAAGGGTAGAGAAACCAAGTAAGATGGTAGGTGGTGGAGGGGCAT 17417

QY 149 CAGAGGGCAGACATACTGAAACCATACACGACGAGAAACCTAGTCAATCTTAATCAACATAGG 208
Db 17418 CAGAGGGCAGACACACCAAAACCAATCAACAGAAACCTAGTCAATCTGATCAGACAC--GG 17475

QY 209 ACCACAGCCTTGCTCTAACTCAATGAACCTAGCCATGCC-CGTGGGCGAACCCAGATGG 267
Db 17476 ACCACAGCCTTGCTCTAACTCAATGAACCTAGCCATGCCACGTGGGGCCACCCAGACAG 17535

QY 268 CGAGTCTATGGTGGAGAGATCTGACAGAAATGTGTCCACTGGAGAAAGGGAAT-GCAAAACC 326
Db 17536 CAGGATCATGACAGAGAGGTCTGACAGAAATGTGTCCACTGGAGAAAGGGAATGGCAACC 17595

QY 327 ACTTCAGTATTCTTGCCTTGGAGACCCCATGAACAGTATGAAAGGCAAAATGATAGAT 386
Db 17596 ACTTCAGTATTCTTGCCTTGGAGACCCCATGAACAGTATGAAAGGCAAAATGATAGAT 17655

QY 387 ACTGAAAGAGAACTCCCCAGGTAGTAGGTGCCCCCATATCTACTGGAGATCAGTGGAG 446
Db 17656 ACTGAAAGAGAACTCCCCAGGTAGTAGGTAGTACCCCATATCTACTGGAGAACTGGAG 17715

QY 447 AAATACTCCAGAAAGAAATGAAGATGAGAGACCGCCAAAGCAAAAGAAATACCCAGGTGG 506
Db 17716 AAATACTCCAGAAAGAAATGAAGTATGAGAGACCGCCAAAGCAAA-----ACCCAGTGTAGA 17770

QY 507 TGTGACTGGTATATAGCAAGGTCCGATGCTGTAAAGAGCAATATTGCATAGAACTGTG 566
Db 17771 TATGACTGGTATAGCAAGGTCCGATGCTGTAAAGAGCAATATTGCATAGAACTGTG 17830

QY 567 GAATGTCCAGTCCATGAATCAAGGCAAAATTTGGAAGTGTCTAAACAGAGATGGCAAGAT 626
Db 17831 GAATGTTCAGTCCATGAATCAAGGCAAAATTTGGAAGTGTCTAAACAGAGATGGCAAGAT 17890

QY 627 GAATGTCAACATTTCTAGAAATCAGCGAACTAAATAGACTGGAATGGTGAATTTAACTC 686
Db 17891 GAACATTCACATTTCTAGAAATCAGCGAACTAAATAGACTGGAATGGTGAATTTAACTC 17950

QY 687 AGATGACCATTTATATCTACTACTGCGGCGAGGAATCCCTCAGAGAAATGGAGTAGCCAT 746
Db 17951 AGATGACCATTTATATCTATTTACTGTGGCAGGAATCCCTTAGTAGAAATGTAGTAGCCAT 18010

QY 747 CATGCTCAACAAAAGAGTCCGAAATGCGATGACTTTGGATGAGTCTCAAAAACGACAGAA 806
Db 18011 CATAGTCAACAAAAGAGTCTCGAAATGCGATGACTTTGGATGAGTCTCAAAAACGACAGAA 18070

QY 807 GATCTCTGTTGTTTCCAAAGCAACCAATTCATATCAGAGTATTCGAAGTCTATGCCCC 866
Db 18071 GATCTCTGTTGTTTCCAAAGCAACCAATTCATATCAGAGTATTCGAAGTCTATGCCCC 18130

QY 867 AACCAAGTAATGCTGAAGAGCTGAAGTTGAACGGTCTCTATGAAGACCTACAAGACCTTTT 926
Db 18131 AACCAAGTAATGCTGAAGAGCTGAAGTTGAAGTGAACGGTCTCTATGAAGACCTACAAGACCTTTT 18190

QY 927 AGAATCAACCCCAAAAGAGTCTCTCTCATTTATAGGGACTGGAATGCAAAAGTAGG 986
Db 18191 AGAATCAACCCCAAAAGAGTCTCTCTCATTTATAGGGACTGGAATGCAAAAGTAGG 18250

QY 987 AAGCAAGAAACACTGGAGTAAACAGGCAAAATTTGGCCTTGGAAATACCGAATGAAGCAGG 1046
Db 18251 AAGTCAAGAAACACTGGGTTAAACAGGCAAAATTTGGCCTTGGAAATACCGAATGAAGCAGG 18310
```

```
QY 1047 GCAAGAGACTAATAGAGTTTTGCCAAGAAATGACCTGGTCTATAGCAACACCTCTTCCA 1106
Db 18311 GCAAGAGACTAATAGAGTTTTGCCAAGAAATGACCTGGTCTATAGCAACACCTCTTCCA 18370

QY 1107 ACAACACAGAGAAGACTCTTACACATGGGACATCACACAGATGGTCAACACCGAAATCAGAT 1166
Db 18371 ACAACACAGAGAAGACTCTTCCACATGGGACATCACACAGATGGTCAACACCGAAATCAGAT 18430

QY 1167 TGATTATATTCTTTGGAGCCAAAGATGGAGAGCTCTATACAGTACGCAAAACCAAGACC 1226
Db 18431 TGATTATATTCTTTGGAGCCAAAGATGGAGAGCTCTATACAGTACGCAAAACCAAGACC 18490

QY 1227 AGGAGCTTACTGTGGCTCAGATCATGAATCCTTTATTGGCAAAATTCAGACTTAAATGAA 1286
Db 18491 AGGAGCTTACTGTGGCTCAGACCATGAATCCTTTATTGGCAAAATTCAGACTTAAATGAA 18550

QY 1287 GAAAGTAGGAAACCACTAGATCACTCAGTGAAGACTTAATTCGAATCCCTTATGATTA 1346
Db 18551 GAAAGTAGAG-AAACCACTAGACCATTCAGGTATAACCTTAAATCAAATCCCTTATGATTA 18609

QY 1347 TACAGTGGAAAGTGAAGAAATAGATTTAAGGCGCTAGATCTGTATAGACAGAGTACTTAATGA 1406
Db 18610 TACAGTGGAAAGTGAAGAAATAGATTTAAGGCGCTAGATCTGTATAGATAGAGTGCCTGATGA 18669

QY 1407 ACTATGGAACAGAGTTTCATGACATTTGTAAGGAGACAGGATCGAGATCCCATGGA 1466
Db 18670 ACTATGGAACAGAGTTTCATGACATTTGTAAGGAGACAGGATCGAGATCCCATGGA 18729

QY 1467 AAAGAAATGCAAAAAAGCAAAATGGCTGTCTGGGAGGCTTACAAATAGCTGTGAAAG 1526
Db 18730 AAAGAAATGCAAAAAAGCAAAATGGATATCTGAGGAGGCGGTACATATAGCTGTGAAAG 18789

QY 1527 AAGAGAAAGTGAAGCAAAAGGAAAGAAAGATAAAGCATCTGAATGCAGAGTTTCCA 1586
Db 18790 AAGAGAAAGTGAAGCAAAAGGAAAGGAGATATAAGCATCTGAATGCAGAGTTTCCA 18849

QY 1587 AA----- 1588
Db 18850 AAGATAGCAAGGAGAGATAAGAAAGCCTTCTCTAGCGATCAGTGCAAAAGATAGAGGA 18909

QY 1589 ----- 1588
Db 18910 AAACATAGATGGGAAAGACTAGAGATCTCTTCAAGAAATTAGAGATACCAAGGAAC 18969

QY 1589 ----- 1588
Db 18970 ATTTTCAAAGATGGGCTCAATAAAGGACAGAAATGGTATGGACCTAACAGAGAGAGAG 19029

QY 1589 ----- 1588
Db 19030 ATATAAGAGAGGTGGCAAGAGTACATNGGAAGAACTGTACAAAAAAGAGCTTCAACAACC 19089

QY 1589 ----- 1588
Db 19090 AGATAATCAGATGGTGTGATCACTGACCTAGAGTACAGACATCCTTGGAAATGTGAAGTCAA 19149

QY 1589 ----- 1588
Db 19150 GTGGGCTTTAGAAAGCATCACTACGAACAAAGCTAGTGGAGGTGATGGAATTCAGTTGA 19209

QY 1589 ----- 1588
Db 19210 GCTATTTCAAATCTTAAAGATGATCCTGTGAAAGTGTGCACTCAATATATGCCAGCAGAT 19269

QY 1589 ----- 1588
Db 19270 TTGGTAACTCAGAGTGGGCCACAGGACTGGAAACGGTCAAGTTTTCATTCCAAATCCCAA 19329

QY 1589 ----- 1588
Db 19330 GAAAGCAATGCCAAGAAATGCCAAACTACCGCACAAAGTGTATTTCATCTCACATGCTAG 19389
```

1589 QY 1596 -----GAATTTC 1596
19390 Db 19449 TAAATAATGCTCAAAATTTCTCAAGCCAGGCTTTCAGCAATACGTGAACCATGAATTC 19449
1597 QY 1656 AGTTGTTCAAGCTGGTTTGAAGAAGTCAGAGAACAGAGACCAAAATTCGCAACATCCT 1656
19450 Db 19509 AGATGTTCAAGCCAGTTTTCGAAGAAGCAGAGAACCAAGATCAAAATTCGCAACATTCG 19509
1657 QY 1716 CTGTATCATGGAAGAACAGAGAGTTCCAGAAAAACATCTATTCTGCTTTATGACTA 1716
19510 Db 19569 CTGGATCATCAAAAAGCAAGAGAGTTCCAAAAAACAATCTATTCTGCTTTATGACTA 19569
1717 QY 1776 TGCAAAAGCCTTTGACTGTGGGGTCAAAATAAAGTGGGAATCTGGAAGGATGGG 1776
19570 Db 19629 TGCCAAAGCCTTTGACTGTGGGATCAAAATAAAGTGGGAATCTGGAAGGATGGG 19629
1777 QY 1836 AATACCAAGCACCCTGACTCTTCAAAAAATTTGATGAGGTGAGGAAGCAACAGT 1836
19630 Db 19689 AATATCAGACCACTGACTTGCCTCTTGAGAAACCTATATCGAGGTGAGGAAGCAACAGT 19689
1837 QY 1896 TAGAAGTGGATGGAACAAACAGAGCTGTTCAAGTAGGAAAAGAGATGTGCAAGCTG 1896
19690 Db 19749 TAGAAGTGGATGGAACAAACAGAGCTGTTCAAGTAGGAAAAGAGATGTGCAATCAATGCTG 19749
1897 QY 1955 TATATTGTCACCGGCTTGTGTTTAACTCTATGCGAG-ACATCATGAGAAACGCTGGCT 1955
19750 Db 19809 TATATTGTCACCGGCTTGTGTTTAACTCTATGCGAGTACATCATGAGAAACGCTGGCT 19809
1956 QY 2015 GGAAGAAGCACAAGCTGGAATCAAGATTGCCGGGAGAAATAGCAATTAACCTCAGATGCG 2015
19810 Db 19869 AGAGGAGCACAAGCTGGAATCAAGATTGCCGGGAGAAATCTCAATAACCTCAGATGCG 19869
2016 QY 2075 AGATGATACCACTTATGCGAGAAAGTGAAGAGAACTAAAGCCTCTGATGAGAGT 2075
19870 Db 19929 AGATGATACCACTTATGCGAGAAAGTGAAGAGAACTAAAGCCTCTGATGAGAGT 19929
2076 QY 2135 GAAAGAGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAAAAGCAAGATCATGGC 2135
19930 Db 19989 GAAAGAGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAAAAGCAAGATCATGGC 19989
2136 QY 2195 ATCTGGTCCATCATCTTCAAGGAATAGATGGGAAACAGTGGAAACAGTGCAGACTT 2195
19990 Db 20049 ATCTAGTCCATCATCTTCAAGGAATAGATGGGAAACAGTGGAAACAGTGCAGACTT 20049
2196 QY 2255 TATTTTGGGGGCTCCAAATCACTGAGATGCTGAGCAATGAGCAATGAAATTAAGA 2255
20050 Db 20096 TATTTTGGGGGCTCCAAATCACTGAGATGCTGAGCAATGAGCAATGAAATTAAGA 20096
2256 QY 2313 CACTTACTCTTTGGAAGAAAGTTA-ACCACCTAGATAGATATTAAGAAAGCAGAGAC 2313
20097 Db 20155 GTTTACTCTTTGGAAGAAAGTTATGACCAAACTAGATAGATATTAAGAAAGCAGAGAC 20155
2314 QY 2373 ATTACTTCCCAAGCCCATCTAGTCAAGGCTATGTTTCCAGTGGTCAATGAT 2373
20156 Db 20215 ATTACTTCCCAAGCCCATCTAGTCAAGGCTATGTTTCCAGTGGTGGTAT 20215
2374 QY 2433 GGATGTGAGAGTTGAGCTGTGAAGAAGCTGAGCACTGAAGAATGATGCTTTGAACTG 2433
20216 Db 20275 GGATGTGAGAGTTGAGCTGTGAAGAAGCTGAGCACTGAAGAATGATGCTTTGAACTG 20275
2434 QY 2493 TGGTGTGAGAGACTCTTGAGTCCCTTGGAGTCAAGAGATCAACCAAGTCCATT 2493
20276 Db 20335 TGGTGTGAGAGACTCTTGAGTCCCTTGGAGTCAAGAGATCAACCAAGTCCATT 20335
2494 QY 2553 CTGAAGGAGTCAAGCCCTGGGATTTCTTTGGAAGGAATGATGCTAAAGCTGAACTCCAG 2553
20336 Db 20395 CTAAGGAGATCAGTCTCTGGTGTTCATTTGAGAGACTGATGCTGAAGTGAATCCAA 20395
2554 QY 2613 TACTTTGGCCACTGATCAAGAGAGTCACTGAGAAAGACCCCTGATGCTGGAGGG 2613
20396 Db 20455 TACTTTGGCCACTGATCAAGAGAGTCACTGAGAAAGACCCCTGATGCTGGAGGG 20455
2614 QY 2673 ATTGGGGGCGAGGAGAGGGGACGACAGAGGATGAGATGGCTGGATGCTACTGAC 2673

20456 Db 20515 ATTGGGGGCGAGGAGAGGGGACGACAGAGATGAGATGGCTGGATGCTACTCCAGAC 20515
2674 QY 2733 TCGATGGAGCTGAGTCTGGTGAATCTCTGGAGTTGGTATGAGCAGGAGGCTGTCT 2733
20516 Db 20575 TCAATGGACATGAGTTTGAATAAATCTGGGAGTTGGTAAATGAACAGGAGGCTGGCGT 20575
2734 QY 2793 GCGCGATTCATGGGTCACAAAGAGTTGGACACGACTGAGCAACTGAACCTGAAC 2793
20576 Db 20635 GCTGTGATTCCTGGGTCACAAAGAGTCAGACACGACTGAGCAACTGAACCTGAAC 20635
2794 QY 2806 TGTACTGAAACCT 2806
20636 Db 20648 TGAACCTGAGCAT 20648
AC092727 163604 bp DNA linear MAM 10-OCT-2002
Bos taurus clone RP42-351K5, complete sequence.
AC092727
AC092727.2 GI:17386265
HIG.
Bos taurus (cow)
Bos taurus
Bos taurus
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 163604)
Akher, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Ho, S.-L., Idol, J.R., Karlins, E., Larc, P., Lee-Lin, S.-Q.,
Legaspi, R., Maduro, Q.L., Maduro, V.B., Masello, C., Mastrian, S.D.,
McCloskey, J.C., McDowell, J., Pearson, R., Prasad, A., Shevchenko, Y.,
Stancic, P., Thomas, J.W., Thomas, P.J., Touchman, J.W.,
Tsaur, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Zhang, L.-H. and Green, E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 163604)
Green, E.D.
Direct Submission
Submitted (21-JUL-2001) NIH Intramural Sequencing Center, 8717
Government Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 163604)
Green, E.D.
Direct Submission
Submitted (06-DEC-2001) NIH Intramural Sequencing Center, 8717
Government Circle, Gaithersburg, MD 20877, USA
4 (bases 1 to 163604)
Green, E.D.
Direct Submission
Submitted (10-OCT-2002) NIH Intramural Sequencing Center, 8717
Government Circle, Gaithersburg, MD 20877, USA
On Dec 6, 2001 this sequence version replaced gi:14993717.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoonhgri.nih.gov
----- Project Information
Center project name: cjb
Center clone name: 351K05

This sequence was finished as follows unless otherwise noted:
all regions were double-stranded, sequenced with an
alternate chemistry, or covered by high quality data
(i.e., phred quality >= 30); an attempt was made to resolve
all sequencing problems, such as compressions and repeats;
all regions were covered by at least one plasmid subclone
or more than one M13 subclone; and the assembly was confirmed
by restriction digest.

CLONE LENGTH: This sequence represents the entire insert of this clone unless otherwise noted. If there are overlapping clones, the overlaps are noted in the beginning and end of the Features section.

FEATURES

```

source
  1. .163604
    /organism="Bos taurus"
    /mol_type="genomic DNA"
    /db_xref="taxon:9913"
    /clone_lib="RP42-351K5"
    /clone_lib="RP42"
    /notes="low quality single stranded/single chemistry region"
  35022..35246
    /note="single clone coverage"
  35403..35706
    /note="single clone coverage"
  13788..137819
    /note="single clone coverage"
  154637..154640
    /note="low quality single stranded/single chemistry region"
  154649..154669
    /note="low quality single stranded/single chemistry region"
  154675..154683
    /note="low quality single stranded/single chemistry region"

Query Match      54.1%; Score 1931.7; DB 1; Length 163604;
Best Local Similarity 75.9%; Pred. No. 0;
Matches 2468; Conservative 0; Mismatches 153; Indels 629; Gaps 13;

QY 162 TACTGAAACCATACACGAGAACTAGTCAATCTATCATCATCTAGACACACAGCCTTGT 221
DB 42552 TAGTGCCCGAGTCACGCCAAAATACTAGTCAATCTATCATCTAGACACACAGCCTTGT 42493

QY 222 CTAACTCAATGAATAAGCCATGCCGTCGGGCAACCCCAAGATGGCGAGGTCAATGGTG 281
DB 42492 CTAACTCAATGAATAAGCCATGCCGTCGGGCAACCCCAAGATGGCGAGGTCAATGGTG 42433

QY 282 AGAGATCTGACAGAAATGTGTCCTAGGAGAGGAAT-GCAACCACTTTCAGTATCTTT 340
DB 42432 AGAGGTCTGACAGAAATGTGTCCTAGGAGAGGAATGCAACCACTTTCAGTATCTTT 42373

QY 341 GCCTTGAGAACCCCATGAACAGATGATAAGAGGCAAAATGATAGGATCTGAAAGAGGAAC 400
DB 42372 GCCTTGAGAACCCCATGAACAGATGATAAGAGGCAAAATGATAGGATCTGAAAGAGGAAC 42313

QY 401 TCCCGAGTCTAGTAGTGCCTCCATATGCTACTGAGATCAGTGGAGAAATTAATCCAGAA 460
DB 42312 TCCCGAGTCTAGTAGTGCCTCCATATGCTACTGAGATCAGTGGAGAAATTAATCCAGAA 42253

QY 461 AGAATGAAGAGATGGAGCAAGCAAAAGAAATACCCAGCTGTGGATGTCAGTGGTGATA 520
DB 42252 AGAATGAAGAGATGGAGCAAGCAAAAGAAATACCCAGCTGTGGATGTCAGTGGTGATA 42193

QY 521 TAAGCAAGGTCGATGCTGTAAAGAGCAATATTGCATAGGAACCTGGAATGTGAGTCCA 580
DB 42192 TAAGCAAGGTCGATGCTGTAAAGAGCAATATTGCATAGGAACCTGGAATGTGAGTCCA 42133

QY 581 TGAATCAAGGCAAAATGGAGGTGTCACACAGAGATGGCAAGTGAATGTCAACATTC 640
DB 42132 TGAATCAAGGCAAAATGGAGGTGTCACACAGAGATGGCAAGTGAATGTCAACATTC 42073

QY 641 TAGGAATCAGCGAACTAAATGGACTGGAATGGGTGAATTTAACTCAGATGACCAATATA 700
DB 42072 TAGGAATCAGCGAACTAAATGGACTGGAATGGGTGAATTTAACTCAGATGACCAATATA 42013

QY 701 TCTACTACTGCGGCGAGGAATCCCTCTCAGAGAAATGGA--GTAGCCATCATGGTCAACAA 758
  
```

```

DB 42012 TCTACTACTGCGGCGAGGAATCCCTCAGAGAAACGGACTGTAGCCATCGTGTCAACAA 41953
QY 759 FAGGTCCGAAATGAGTACTTGGATGAGTCTCAAAACGACAGAAATGATCTCTGTTTG 818
DB 41952 AAGAGTCTGAAATGAGTACTTGGATGCAATCTCAAAACGACAGAAATGATCTCTGTTTG 41893
QY 819 TTTCAGGCAAAACCAATTCATATACAGTAAATCAAGTCTATGCCCCAACAGGATGTC 878
DB 41892 TTGCCAAGGCAAAACCAATTCATATACAGTAAATCAAGTCTATGCCCCAACAGGATGTC 41833
QY 879 TGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 938
DB 41832 TGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 41773
QY 939 CAAAAAAGATGCTCTTCTCATATAGGGGACTGGAATGCAAAAGTGAAGGCAAAAGAAAC 998
DB 41772 CAAAAAAGATGCTCTTCTCATATAGGGGACTGGAATGCAAAAGTGAAGGCAAAAGAAAC 41713
QY 999 ACCTGAGTGAAGCAAAATTTGGCTTTGGAATACGGAATGAAGCAGGCAAGACTAAT 1058
DB 41712 ACCTGAGTGAAGCAAAATTTGGCTTTGGAATACGGAATGAAGCAGGCAAGACTAAT 41653
QY 1059 AGAGTTTTGCCAAGAAATGCACTGTCATAGCAAAACCCCTCTTCCAAACACACAGAG 1118
DB 41652 AGAGTTTTGCCAAGAAATGCACTGTCATAGCAAAATGCACTCTTCCAAACACACAGAG 41593
QY 1119 AAGACTCTACATCGACATCACAGATGCCAGATGGTCAACACCGAAATCAGATGATATTTCT 1178
DB 41592 AAGA--CTACACATGGACATCACAGATGGTCAACACCGAAATCAGATGATATTTCT 41535
QY 1179 TTGCAGCCAAAGATGGAGAGCTCTATACAGTCAAGCAAAACCAAGCAGGAGCTTACTG 1238
DB 41534 TTGCAGCCAAAGATGGAGAGCTCTATACAGTCAAGCAAAACCAAGCAGGAGCTTACTG 41475
QY 1239 TGCTCAGATCATGAATCCCTTTATTTGCCAAATTCAGACTTAAATTTGAAGAAATGAGGAA 1298
DB 41474 TGCTCAGATCATGAATCCCTTTATTTACCAATTCAGACTTAAATTTGAAGAAATGAGGAA 41415
QY 1299 AACCACTAGATCACTCAGTGAAGACTTAATCCATCCCTTATGATTTATACAGTGAAGT 1358
DB 41414 AACCACTAGATCACTCAGTGAAGACTTAATCCATCCCTTATGATTTATACAGTGAAGT 41355
QY 1359 GAGAAATAGATTTAAGGCGCTAGATCTGATAGACAGAGTACCTAATGAATGATGACAGA 1418
DB 41354 GAGAAATAGATTTAAG--CCTAGATCTGATAGTAGAGTGCTCTGATGAATGAAGA 41296
QY 1419 GGTTCATGATGATGACAGGAGACAGGATCGAGACCACTCCCATGGAAGAAATGCA 1478
DB 41295 GGTTCATGATGATGACAGGAGACAGGATCGAGACCACTCCCATGGAAGAAATGCA 41236
QY 1479 AAAAGCAAAATGCTGTCTGGGGAGGCTTACAAATAGCTGTGAAAAGAGAGAGAGTGA 1538
DB 41235 AAAAGCAAAATGCTGTCTGGGGAGGCTTACAAATAGCTGTGAAAAGAGAGAGTGA 41179
QY 1539 AAGCAAGGAAAAAGGAAAGATAAAGCATCTGAATCGAGAGTTCGAAA----- 1588
DB 41178 AAGCAAGGAAAAAGGAAAGATAAAGCATCTGAATCGAGAGTTCGAAAAGAGAGTGA 41119
QY 1589 ----- 1588
DB 41118 AAGAGATAAAGAGCCTTCTTCAGCTATCAATGCAAAAGAACTAGAGGAAAAACAGAGAT 41059
QY 1589 ----- 1588
DB 41058 GGAAGAGACTAGAGATCTCTTCAAGAAATTAGAGTACCAAGGGAACATTTTCATGCAAA 40999
QY 1589 ----- 1588
DB 40998 GATGGCTCCATAAAGGACAGAAATGGTATGACCTTAACAGAGCAGAGATATTAAGAA 40939
QY 1589 ----- 1588
DB 40938 GAGGTGGCAAGAAATACCCGAAAGAACTGTACAAAAGAAATCTTCACATCCAGATAATCA 40879
  
```

QY 1589 ----- 1588
 Db 40878 TGATGGTGTGATCACTGACCTAGAGCCAGACATCCTGGAATGTAAAGTCAAGTGCAGCCTT 40819
 QY 1589 ----- 1588
 Db 40818 AGAAGCATCACTATGACAAAGCTAGTGGAGGTGATGGAATCCAGTGGAGCTATTTC 40759
 QY 1589 ----- 1588
 Db 40758 AATCCTGAAGATGATGCTGTGAAGTGTCTGCACTCAATATGCCAGCAAAATTTGGAAGAAC 40699
 QY 1589 ----- 1588
 Db 40698 TCAGCAGTGGCCACAGACTAGAAAAGGTGAGTTTTCATCTGATGCCAAGAAAGGCAA 40639
 QY 1589 ----- 1588
 Db 40638 TGCCAAAGATGCTCAAACTACCAACAGAGTTGCCACCCATCTCACACGCTAGTAAAGTAAT 40579
 QY 1589 ----- 1588
 Db 40578 GCTCAAAATCTCAAGCCAGCTTCAGCAATATGTGAATGGAATTCCTGATGTTCAA 1506
 QY 1607 GCTGGTTTTAGAAAAGTCCAGAGGAACAGAGACCAAAATTTGCCAACATCTCTGTATCATG 1666
 Db 40518 GCTGGTTTTAGAAAAGGACAGAGGAACAGAGATCAATTTGCCAACATCTCTGTATCATG 40459
 QY 1667 GAAAAAGCAAGAGATTTCCAGAAAAACATCTATTTCTGCTTTATGACTATGCAAAAGCC 1726
 Db 40458 GAAAAAGCAAGAGATTTCCAGAAAAACATCTATTTCTGTTTATGACTATGCAAAAGCC 40399
 QY 1727 TTTGACTGTGGGGTCAAAATAAATCTGGAATAATCTGAAAGGATGGGAATACACAGAC 1786
 Db 40398 TTTGACTGTGGATCAAAATACCAATACCTGGAATAATCTTCAAGAGATGGGAATACACAGAC 40339
 QY 1787 CACCTGACCTGACTCTTGAAAAATTTGTATGACGATGAGGAACCAAGTGAAGTGA 1846
 Db 40338 CACCTGACCTGCTCTTGAAAAATCTGTATGACGATGAGGAACCAAGTGAAGTGA 40279
 QY 1847 CATGGAACACAGACTGTTCCAAAGTGAAGGATGATGCAAGGCTCTATATTTGCTCA 1906
 Db 40278 CATGGAACACAGACTGTTCCAAAGTGAAGGATGATGCAAGGCTCTATATTTGCTCA 40219
 QY 1907 CCCGGCTTGTTTAACTTCTATGACAGAG-ACATCATGAGAAACGCTGGGCTGGAAGAGCA 1965
 Db 40218 CCCTGCTTATTTAACTTCTATGACAGAGTACATCATGAGAAATGCTGGACTAGAAGAAACA 40159
 QY 1966 CAAGCTGGATCAGATGTCGGGAGAAATAGCAATACCTCAGATGACAGATGATACC 2025
 Db 40158 CAAGCTGGAATCAGATGTCGGGAGAAATATCAATACCTCAGATGACAGATGACACC 40099
 QY 2026 ACCCTTATGGCAGAAAGTGAAGAGAACTAAAAAGCCTCTTGATGAAGGTGGAAGAGGAG 2085
 Db 40098 ACCCTTATGGCAGAAAGTGAAGAGAACTAAAAAGCCTCTTGATGAAGGTGAAGTGGAG 40039
 QY 2086 AGTGAAGAAAGTGGCTTAAGCTCAACATTCAGAAAACGAAGATC-----ATGCGATCTGG 2141
 Db 40038 AGTGAAGAAATTTGGCTTAAAGCTCAACATTCAGAAAACGAAGATCATTGGATGGATCCAG 39979
 QY 2142 TCCCATCACTTCATGGGAATAGATGGGAAACAGTGGAAACAGTGCAGACTTTATTTT 2201
 Db 39978 TCCCATCACTTCATGGGAATAG-----GGAAACAGTGCAGAAATTTATTTT 39932
 QY 2202 TTGGGGGGCTCCAAAATCACTGACAGATGGTGAATGACGCAATGAAATTAAGAAACACTTA 2261
 Db 39931 TT--GGGTCTCCAAAATCACTGACAGATGGTGAATGACGCAATGAAATTAAGAAACACTTA 39878
 QY 2262 CTCCTTCGAGAAAAGTTA--ACCAACCTAGATAGCATATTTGAAACAGACAGACATTAACC 2319
 Db 39877 CTCCTTCGAGGCAAGTTTATGTCCACGTAGATAGCATATTCAGAACAGACAGATTAAT 39818

QY 2320 TTGCCAACCAAGCCCATCTAGTCAAGGCTATGGTTTTTCCAGTGGTCACTGTATGGATGT 2379
 Db 39817 TTGCCAACCAAGGTCGCTTAGTCAAGGCTATGGTTTTTCCAGTGGTCACTGTATGGATGT 39758
 QY 2380 GAGAGTTGCACTGTGAAGAAAGCTGAGCACCTGAAGAAATGTGCTTTTGAACCTGTGTGTGT 2439
 Db 39757 GAGAATTTGAGTGTGAAGAAAGCTGAGTGTCTGAAGAAATGTGCTTTTGAACCTGTGTGTGT 39698
 QY 2440 TGGAGAAAGCTCTTGAGAGTCCCTTGGACTGCAAGAGAGATCCAAACAGTCCATTTCTGAAG 2499
 Db 39697 TGGAGAAAGCTCTTGAGAGTCCCTTGGACTGCAAGAGAGATCCAAACAGTCCATTTCTGAAG 39638
 QY 2500 GAGATCAGCCCTGGGATTTCTTTGGAAGGAATGATGCTAAAGCTGAAACTCCAGTACTTT 2559
 Db 39637 GAGATCAGCCCTGGGATTTCTTTGGAAGGAATGATGCTAAAGCTGAAACTCCAGTACTTT 39578
 QY 2560 GGCCACCTGATCAGAAAGCTGACTCACTGGAAGAAAGCCCTGATGCTGGAGGGAATGGG 2619
 Db 39577 GGCTCTCTCATCGAAGAGTTGACTCATTTGGAAGAAAGCTGATGCTGGAGGGAATGGG 39518
 QY 2620 GGCAGAGAGAGAGGAGGACGACAGAGGATGAGATGCTGGATGGCATCACTGACTCGATG 2679
 Db 39517 GGCAGAGAGAGAGGAGGACGACAGAGGATGAGATGCTGGATGGCATCACTGACTCGATG 39458
 QY 2680 GACGTGAGTCTGGGTGAACTCCTGAGTTGATGATGAGGAGGCTGCTGCTGCGCG 2739
 Db 39457 GATGTGAGTCTGAGTGAACCTCGGAGTTGATGATGAGGAGGCTGCTGCTGCGCG 39398
 QY 2740 ATTCTATGGGTCAAAAGAGTTGGACAGACTGAG--CAACTGAACTGAACTGAACTGTA 2797
 Db 39397 ATTCTATGGGTGCAAAAGAGTTGGACAGACTGAGTTGAGCTCACTGAACTGAACTGTA 39338
 QY 2798 CTGAAACCTT 2807
 Db 39337 CTTTGGCAT 39328

RESULT 32

AC092410

LOCUS

DEFINITION

AC092410.1

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AC092410 168497 bp DNA linear HTG 04-JUL-2001
 Bos taurus clone RP42-245C6, WORKING DRAFT SEQUENCE, 5 unordered pieces

AC092410.1 GI:14595784

HTG; HTGS_PHASE1; HTGS_DRAFT.

Bos taurus (cow)

Bos taurus

Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovidae; Bovinae; Bos.

1 (bases 1 to 168497)

Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W.,

Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L.,

Granite, S., Guan, X., Gupta, J., Ho, S.-L., Idol, J.R., Karlins, E.,

Lee-Lin, S.-Q., Legaspi, R., Lim, M., Maduro, O.L., Maduro, V.B.,

Masello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J.,

Pearson, R., Prasad, A., Shevchenko, Y., Snyder, B., Stantropop, S.,

Thomas, J.W., Thomas, P.J., Touchman, J.W., Tsurgon, C., Vogt, J.L.,

Walker, M.A., Wetherby, K.D., Zhang, L.-H. and Green, E.D.

NISC Comparative Sequencing Initiative

Unpublished

2 (bases 1 to 168497)

Green, E.D.

Direct Submission

Submitted (04-JUL-2001) NIH Intramural Sequencing Center, 8717

Grovermont Circle, Gaithersburg, MD 20877, USA

----- Genome Center

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: http://www.nisc.nih.gov

Contact: nisc.mouse@nih.gov

----- Project Information

----- Project name: cjc

```

Center clone name: 245C06
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 165686 bases at least Q40
Consensus quality: 167007 bases at least Q30
Consensus quality: 167085 bases at least Q20
Insert size: 159000; agarose-fp
Insert coverage: 12.17x in Q20 bases; agarose-fp
Quality coverage: 11.52x in Q20 bases; sum-of-coverage
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 12474: contig of 12474 bp in length
* 12475: gap of unknown length
* 12575: contig of 15003 bp in length
* 27577: gap of unknown length
* 27578: contig of 18860 bp in length
* 27678: gap of unknown length
* 46337: gap of unknown length
* 46338: contig of 33766 bp in length
* 80403: gap of unknown length
* 80504: contig of 87994 bp in length.
*
FEATURES
      source
      1..168497
        /organism="Bos taurus"
        /mol_type="genomic DNA"
        /db_xref="taxon:9913"
        /clone="RP42-245C6"
        /clone_lib="RP42"
        1..12474
          /note="assembly_fragment"
          12575..27577
            /note="assembly_fragment"
            clone_end:7
            vector_side:left
            27678..46337
              /note="assembly_fragment"
              46338..80403
                /note="assembly_fragment"
                80504..168497
                  /note="assembly_fragment"
                  clone_end:SP6
                  vector_side:right
Query Match      54.0%; Score 1930.2; DB 1; Length 168497;
Best Local Similarity 76.2%; Pred. No. 0;
Matches 2376; Conservative 0; Mismatches 143; Indels 600; Gaps 6;

QY 276 TGGTGGAGAGATCTGACAGAAATGTGGTCCACTGAGAGAGGGAAT-GCAAAACCACTTCAGT 334
DB 143498 TGGGGTTGACACTTGACAGATGTGGTCCACTGAGAGAGGGAATGGCAAAACCACTTCAGT 143557
QY 335 ATTCTTGCTTGAAGAACCCCATGAACTATGAAAGCCAAATGATAGGACTGAAAG 394
DB 143558 ATTCTTGCTTGAAGAACCCCATGAACTATGAAAGCCAAATGATAGGACTGAAAG 143617
QY 395 AGGAACCTCCAGGTCAGTAGTGCCCATATGCTTACTGGAGATCAGTGGAGAAATAACT 454
DB 143618 AGAACTCCCGAGTCAGTAGTGCCCATATGCTTACTGGAGATCAGTGGAGAAATAACT 143677
QY 455 CCAGAAAGATCAAGAGATGGAGCCAAAGCAAAAGAAATACCCAGCTGTGGATGTGACTG 514
DB 143678 CCAGAAAGATCAAGGGATGGAGCCAAAGCAAAAGCAATATCCAGCTGTGGATGTGACTG 143737
QY 515 GTGATATAAGCAGGTCGATGCTGTAAAGAGCAATATTTGCATAGGAACCTGGAAATGTCA 574

Db 143738 GTGATAGAGCAGGTCGAATGCTGTAAAGAGCAATATTTGCATAGGAACCTGGAAATGTCA 143797
QY 575 GGTCCATGAATCAAGGCAAAATTTGAAGTGTCAAAAGAGATGGCAAGAGTGAATGTCA 634
DB 143798 GGTCCATGAATCAAGGCAAAATTTGAAGTGTCAAAAGAGATGGCAAGAGTGAATGTCA 143857
QY 635 ACATCTAGGAATCAGGCAAAATTTGAAGTGTCAAAAGAGTGGTGAATTTAACTCAGATGACC 694
DB 143858 ACATCTAGGAATCAGGCAAAATTTGAAGTGTCAAAAGAGTGGTGAATTTAACTCAGATGACC 143917
QY 695 ATTATATCTACTACTGCGGGCAGGAATCCCTCAGAGAAATGGAGTAGCCATATGGTCA 754
DB 143918 ATTATATCTACTACTGCGGGCAGGAATCCCTCAGAGAAATGGAGTAGCCATATGGTCA 143977
QY 755 ACAAGAGTCCGAAATTCGAGTGTCAAAAGAGTGGTGAATTTAACTCAGATGACTCTG 814
DB 143978 ACAAGAGTCCGAAATTCGAGTGTCAAAAGAGTGGTGAATTTAACTCAGATGACTCTG 144037
QY 815 TTTGTTTCCAAAGGCAAAACCAATTCAGAGTGTCAAAAGTGTATGCCCAACCAAGTA 874
DB 144038 TTTGTTTCCAAAGGCAAAACCAATTCAGAGTGTCAAAAGTGTATGCCCAACCAAGTA 144097
QY 875 ATGCTGAAGAGCTGAAGTTGAACGGTCTATGAAGACCTTACAGACCTTTTAGAATAA 934
DB 144098 ACATCTGAAGAGCTGAAGTTGAACGGTCTATGAAGACCTTACAGACCTTTTAGAATAA 144157
QY 935 CACCCAAAAAGATGTCTCTCTCATATAGGGGAGTGGAAATGCAAAAGTAGGAAGCAAG 994
DB 144158 CACCCAAAAAGATGTCTCTCTCATATAGGGGAGTGGAAATGCAAAAGTAGGAAGCAAG 144217
QY 995 AAACACCTGGAGTAACAGGCAAAATTTGGCTTGGAAATCGGAATGAAGCAGGCGCAAGAC 1054
DB 144218 AAACACCTGGAGTAACAGGCAAAATTTGGCTTGGAAATCGGAATGAAGCAGGCGCAAGAC 144277
QY 1055 TAATAGAGTGTGGCAAGAAATGCACTGTGTATAGCAAAACACCTCTTCCAAACACACA 1114
DB 144278 TAATAGAGTGTGGCAAGAAATGCACTGTGTATAGCAAAACACCTCTTCCAAACACACA 144337
QY 1115 AGAGAGAGCTCTACACATGGACATCACCAGATGFTCAACCGGAATCAGATTTGATTATA 1174
DB 144338 AGAGAGAGCTCTATACATGGACATCACCAGATGFTCAACCTGAAATCAGATTTGATTATA 144397
QY 1175 TTCTTTGACGCAAGATGGAGAGCTCTATACAGTGTGCAAAACACAGACAGCAGGAGCTT 1234
DB 144398 TTCTTTGACGCAAGATGGAGAGCTCTATACAGTGTGCAAAACACAGACAGCAGGAGCTT 144457
QY 1235 ACTGTGGCTCAGATCATGAACTCTTATTTGCCAAATTCAGACTTAAATTTGAAGAAAGTAG 1294
DB 144458 ACTGTGGCTCAGACCATGAACTCTTATTTGCCAAATTCAGACTCAGATTTGAAGAAAGTAG 144517
QY 1295 GGAAACCACTAGATCATCAGGTAGAGCTTAATTCGAATCCATCCCTTATGATTTATCAGTGG 1354
DB 144518 GGAAACCACTAGATCATCAGGTAGAGCTTAATTCGAATCCATCCCTTATGATTTATCAGTGG 144577
QY 1355 AAGTGAAGATAGATTTAAGGCTCTATAGCTGTATAGACAGAGTACTAATGAATATATGA 1414
DB 144578 AAGTGAAGATAGATTTAAGGCTCTATAGCTGTATAGACAGAGTACTAATGAATATATGA 144637
QY 1415 CAGAGGTTTCATGACATTTGTACAGGAGACAGGATCCAGACCATCCCTATGAAAAAGAAAT 1474
DB 144638 ATGAGGTTTCATGACATTTGTACAGGAGACAGGATCCAGACCATCCCTATGAAAAAGAAAT 144697
QY 1475 GCAAAAGCAAAATGGCTGTCTGTGGGAGGCTTACAAATAGCTGTGAAAAAGAGAGAG 1534
DB 144698 GCAAAAGCAAAATGGCTGTCTGTGGGAGGCTTACAAATAGCTGTGAAAAAGAGAGAG 144757
QY 1535 TGAAGAGCAAAAGGAAAAAGGAAAGATAAAAGCATCTGAATGCAGAGTTCCAAA----- 1588
DB 144758 CGAAAGCAAGAGAGAGAGAGAGATATATGCACTGTAATGCAGACTTCCAGAGATAG 144817
QY 1589 ----- 1588

```

Db	144818	CAAGGAGAGATAGAAAGCGCTTCTCTCAGCGATCAATTCGAAGAAATAGAGGAAACAACA	144877
Qy	1589	-----	1588
Db	144878	GAATGGGAAGACTAGGAGTCTCTTCAAGAAATCAGAGATACCAAGAAACATTTTCATG	144937
Qy	1589	-----	1588
Db	144938	CANAGATGAGCTCGATATAAGGACAGAAATGGTATGGACCTAAACAGAGCACAAGATATTA	144997
Qy	1589	-----	1588
Db	144998	AGAGAGATGGCAAGAAATACACAGAAGACCTGTACAAAAAGATCTTCAGACCCAGATA	145057
Qy	1589	-----	1588
Db	145058	ATCACGATGGTGTGATCACTCACCTAGAGCCAGACATCTCGAATGTGAAGTCAAGTGGG	145117
Qy	1589	-----	1588
Db	145118	CCTTAAGAGCATCACTACGAACAAAGCTAGTGGAGGTGATGGAATTCAGTTCAGCTAT	145177
Qy	1589	-----	1588
Db	145178	TCCAAATCCTGAAAGATGATGCTGTGAAGTCTCTCACTCAATATGCCAGCAAAATTTGGA	145237
Qy	1589	-----	1588
Db	145238	AAACTCAGACTGGCCACAGACTTGGAAAGGTCAGTTTTTGTTCCAATCCAAAGAAAG	145297
Qy	1589	-----	1588
Db	145298	ACRATGACAAAGATGCTCAAGTACCGTACAAATTGCACTCATCTCACAGGCTAGTAAG	145357
Qy	1589	-----	1601
Db	145358	TAATGCTCAAAATTTCTCCAAGCCAGGCTTCAGCAATATGTGAACCGTGAACTTCAGATG	145417
Qy	1602	TTCAAGCTGGTTTTAGAAAAAGTCAGAGGAACACAGAGACCAAAATTCGCCAACATCTCTCTGA	1661
Db	145418	TTCAAGCTGGTTTTAGAAAAAGTCAGAGGAACACAGAGTCAAAATTCGCCAACATCTGCTGGA	145477
Qy	1662	TCATGAAAAAGCAAGAGAGTTCAGAAAAACATCTATTTCTGCTTTATTGCACTATGCAA	1721
Db	145478	TCATAGAAAAAGCAAGAGAGTTCAGAAAAACATCTCTTTCTGCTTTATTGCACTATGCCA	145537
Qy	1722	AAGCCTTTGACTGTGGGGTCACAATAAATCTGTGGAATAATCTCAAAAGGATCGGAATAC	1781
Db	145538	AAGCCTTTGACTGTGTGGATCACAAATAAATCTGTGGAATAATCTCAAAAGAGATCGGAATAC	145597
Qy	1782	CAGACCACTGACCTGACTCTCTTGAAAAATTTGTATGCAGGTTCAGGAAGCAACAGTTAGAA	1841
Db	145598	CAGACCACTGATCTCTCCTCTTCAGAAATTTGTATGCAGGTTCAGGAAGCAACAGTTAGAA	145657
Qy	1842	CTGCACATGGAACAAACAGACTGGTTCCAGTAGAAAAAGAGTATGTCAAGCTGTATAT	1901
Db	145658	CTGCACATGGAACAAACAGACTGGTTCCAAATAGGANAAGGATTCGTCAAGGCTGTTTAT	145717
Qy	1902	TGTCACCCGCTGTGTTTTAACTTCTATGCAGAG-ACATCATGAGAAACGCTGGGCTGGAAG	1960
Db	145718	TGTCACCCCTGTTTATTATTAATCTTATATGCAGAGTACATCATGAGAAACGCTGGACTGAAG	145777
Qy	1961	AAGCACAAGCTGGAATCAAGATTTCCGGGAGAAATAGCAATTAACCTCAGATATGCAGATG	2020
Db	145778	AAACAACAGCTGGAATCAAGATTTCCAGGAGAAATATCAATAACCTCAGATATGCAGATG	145837
Qy	2021	ATACCACCCCTTATGGCGAAGAGTCAAGAGAACTTAAAAAGCCTCTTCTCATGAAGTGAAG	2080
Db	145838	ACACCACCCCTTATGGCGAAGAGTCAAGAGAACTCAAAAGCCTCTTCTCATGAAGTGAAG	145897
Qy	2081	AGAGAGTGAAGAGTTGGCTTAAAGCTCAACATTTCAAAAAACGAAGATCATGGCATCTG	2140
Db	145898	TAGAGAGTGAAGAGTTGGCTTAAAGCTTAAATTCAGAAAAACGAAGATCATGGCATCCG	145957

QY	2141	GTCCCATCATTCTGCGAAATAGATGGGAAACAGTGGAAACAGTGTGACACTTTATT	2200
Db	145958	GTCCACCACCTTCATGGGAAATAGATGGGAAACAGTGGAAACAGTGTGACACTTTATT	146017
QY	2201	TTTGGGGGCTCCAAAATCACTGCAGATGGTGACTGCGAGCCATGAATTAAGAACACTT	2260
Db	146018	TT--CTGGGCTCCAAAATCACTGCAGATGGTGACTGCGAGCCATGAAGTAAAGAACACTT	146075
QY	2261	ACTCCTTGGAAAGAAAGTTA--ACCAACTAGATAGCATATTGAAAGCAGAGACATTAC	2318
Db	146076	ACTCCTTGGAGGGAAGTTATGACCAACTAGATAGCATATTCAAAGCAGAGACATTAC	146135
QY	2319	CTTGCCAAACAAGCCCATCTAGTCAAGGCTATGGTTTTTCCAGTGGTCATGATFGGATG	2378
Db	146136	TTTGCCAAACAAGATCCATCTAAATCAAGGCTATGGTCTTCTGTGGTCATGATGGATG	146195
QY	2379	TGAGATTGGACTGTGAAGAAAGCTGAGCACTGAAGANTTGATGCTTTTGAAGTGTGGTG	2438
Db	146196	TGAGATTGGACTGTGAAGAAAGGCTGAGCGCTGAAGAATTGATGCTTTTGAAGTGTGGTG	146255
QY	2439	TTGAGAAAGACTCTTGAGAGTCCCTTGGAGCTGCAAGGAGATCCCAACCACTCATTTCTGAA	2498
Db	146256	TTGAGAAAGACTCTTGAGAGTCCCTTGGAGCTGCAAGGAGATCCCAACCACTCATTTCTGAA	146315
QY	2499	GGAGATCAGCCCTGGGATTTCTTTGGAAAGGAATGATGCTAAAGCTGAAACTCCAGTACTT	2558
Db	146316	AGAGATCAGCCCTGGGATTTCTTTGGAAAGGAATGATGCTAAAGCTGAAACTCCAGTACTT	146375
QY	2559	TGGCCACTGATCAGAAAGAGCTGACTCACTGGAAAAGACCCTGATGCTGGGAGGGATTGG	2618
Db	146376	TGGCCACTGATCAGAAAGAGTTGACTTATGGAAAAGACTCTGATGCTGGGAGGGATTGG	146435
QY	2619	GGGCAGGAGGAGAGGGGACGACAGAGGATGAGATGGCTGGATGGCATCACTGACTCGAT	2678
Db	146436	GGGCAGGAGGAGAGGGGACGACAGA--GATGAGATGGCTGGATGGATCACTGACTCTAT	146494
QY	2679	GGACGTGAGTCTGGGTGAACTCTCTGGAGTTGGTGATGACAGGAGGCGCTCTCTGGGCG	2738
Db	146495	GGACGTGAGTCTGAGTGAACTCCGGGAGTTGGTGATGACAGGAGGCGCTGGCATGCTGT	146554
QY	2739	GATTTCATGGGGTCACAAAGAGTTGGACACGACTGAGCAACTGAACTGAACTGTA	2797
Db	146555	GATTTCATGGGGTTGCCAAAGAGTCAGACATGACTGAGTCTGATCTGATGATGATA	146613
RESULT 33			
AC130787		173952 bp	linear
LOCUS		Bos taurus clone RP42-205E24,	HTG 27-SEP-2002
DEFINITION		WORKING DRAFT SEQUENCE, 3 ordered	
ACCSSION		pieces:	
VERSION		AC130787	
KEYWORDS		AC130787.2	
		HTG; HTGS PHASE2; HTGS DRAFT.	

RESULT 33	TITLE
AC130787	JOURNAL
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	

REFERENCE 2 (bases 1 to 173952)
AUTHORS Green,E.D.
TITLE Direct Submission
JOURNAL Submitted (14-AUG-2002) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
REFERENCE 3 (bases 1 to 173952)
AUTHORS Green,E.D.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-2002) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
COMMENT On Sep 27, 2002 this sequence version replaced gi:22218457.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: <http://www.nisc.nih.gov>
Contact: nisc_zoo@nhgri.nih.gov
----- Project Information
Center project name: ddb
Center clone name: 206E24

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig, has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 171764 bases at least Q40
Consensus quality: 172948 bases at least Q30
Consensus quality: 173503 bases at least Q20
Insert size: 176000; agarose-fp
Insert size: 173752; sum-of-contigs
Quality coverage: 5.74x in Q20 bases; agarose-fp
Quality coverage: 5.82x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 89789: contig of 89789 bp in length
* 89790 89889: gap of unknown length
* 89890 96759: contig of 8870 bp in length
* 96850 96859: gap of unknown length
* 96860 173952: contig of 77093 bp in length.
Location/Qualifiers
1..173952
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="RP42-206E24"
/clone_lib="RP42"
1..89789
/note="assembly_fragment
clone end.T7
vector_side:left"
89890..96759
/note="assembly_fragment"
96850..173952
/note="assembly_fragment"

FEATURES
source
1..173952
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="RP42-206E24"
/clone_lib="RP42"
1..89789
/note="assembly_fragment
clone end.T7
vector_side:left"
89890..96759
/note="assembly_fragment"
96850..173952
/note="assembly_fragment"

missing SP6 clone_end from 3' end of insert"	
Query Match	53.7%; Score 1919.1; DB 1; Length 173952;
Best Local Similarity	74.5%; Pred. No. 0;
Matches 2536; Conservative	0; Mismatches 239; Indels 627; Gaps 16;
QY	1 GTCAGGAAGGGCGGCGAGTGGAGGAGGTACCTACCTCGTCCAAAGGTAAGGAGCAGTAGC 60
DB	99257 GTCAGAGGGGCGAGCTGTGAGGAGA----TACTGCTCGTCCAAAGTGTAGGAGCAGCAGC 99311
QY	61 TCGCGTTTCTGGAGAGCCCTAAAGAGATACCCCAAGGCCCAAGGTAAAGAGAAACCCCAAG 120
DB	99312 TCGCGTTTCTGGAGAGCCCAATGAAGAGACACCCCAAGTCCAAAGGTAAAGAGAAACCCCAAG 99371
QY	121 TAAAGATGTAGGTGTTGTGAGAGGGCATCAGAGGGCAGACATACCTAGTAAACCATACAGCCA 180
DB	99372 TAAAGATGTAGGTGTTGTGAGA--GGCATCAGAGGGCAGACAC--AAACCATATACACA 99428
QY	181 GAAACTAGTCAATCTAATCATCACTAGGACACAGCCCTTGTCTTAATCAATGAACCTAAG 240
DB	99429 GAAACTAGTCAATCTAATCATCACTAGGACACAGCCCTTGTCTTAATCTAGTGAACCTAAG 99488
QY	241 CCATGCC--CGTGGGGCAACCCAAAGATGGGCGAGGTTCATGGTGGAGAGATCTGACAGAAATG 298
DB	99489 CCATGCCATGTGGGGCCCAACCAAGTCAGCCGGTTCACGGTGGAGAGGTCTGACAGAAATG 99548
QY	299 TGGTCCACTGGAGAGGG--AATGCAAAACCACTTCAGTATTTCTTGCCTTGAGAACCCCATG 357
DB	99549 TGGTCCACTGGAGAGGGTAAAGGCAACCACTTCAGTATTTCTTGCCTTGAGAACCCCATG 99608
QY	358 AACAGTATGAAAGCAAAATGATAGGATCTGAAGAGGAACCTCCCAAGTTCAGTAGGT 417
DB	99609 AACAGTATGAAAGCAAAATGATAGGATCTGAAGAGGAACCTCCCAAGTTCAGTAGGT 99668
QY	418 GCCCATATGCTACTGGAGATCAGTGGAGAAATACTCCAGAAAGAAATGAAGAGATGGAG 477
DB	99669 GCCCATATGCTACTGGAGATCAGTGGAGAAATACTCCAGAGAGATGAAGGGATGGAG 99728
QY	478 CCAAGAGCAAAAGAAATACCCAGCTGTGGATGTGATGTGATTAAGCAAGTCCGATGC 537
DB	99729 CCAAGAGTGAACCAATACCCAGCTGTGGATGTGATGTGATTAAGCAAGTCCGATGC 99788
QY	538 TCTAAGAGCAATATTGCATAGGAACCTGGATGTCAGGTCCATGAATCAAGGCAAAATG 597
DB	99789 TCTAAGAGCAATATTGCATAGGAACCTGGATGTCAGGTCCATGAATCAAGGCAAAATG 99848
QY	598 GAAAGTGGTCAACAAGA--GATGGCAAGAGTGAATGTCAACATTTAGGAATCAGCGAAT 656
DB	99849 GAAATAGTCAACAAGAGAGTGGCAAGAGTGAACGTCGACATTTCTAGGAATCAGCGAAT 99908
QY	657 AAAATGGAGTGAATGGGTGAATTTAACTCAGATGACCATTTATATCTACTACTGCGGCA 716
DB	99909 AACATGGAGTGAATGGGTGAATTTAACTCAGATGACCATTTATATCTACTACTGCGGCA 99968
QY	717 GGAATCCCTCTAGAAGAAATGGAGTAGCCCATCATGTCACCAAAAGAGTCCGAAATGCAAT 776
DB	99969 GGAATCCCTCTAGAAGAAATGGAGTAGCCCATCATGTCACCAAAAGAGTCCGAAATGCAAT 100028
QY	777 ACTTGGATGACGTCTCAAAACGACAGATGATCTCTGTTTTCGAGGCAAAACCATTT 836
DB	100029 ACTTGGATGACATCTCAAAATGACAGATGATCTCTGTTTTCGAGGCAAAACCATTT 100088
QY	837 CAATATCAGTAATCCCAAGTCTATGCCCAACCAAGTAAATGCTGAAGAGAGCTGAAGTTGA 896
DB	100089 CAATATCAGTAATCCCAAGTCTATGCCCAACCAAGTAAATGCTGAAGAGAGCTGAAGTTGA 100148
QY	897 ACGGTCTTATGAAGACCTTACAAGACCTTTTGAAGCTAAACCCCAAAAGATGTCCTCT 956
DB	100149 ATGGTTCTTATGAAGACCTTACAAGACCTTTTGAAGCTAAACCCCAAAAGATGTCCTCT 100208
QY	957 CATTATAGGGGAGCTGAATGCAAAAGTAGGAGCAAGAAACACCTGGAGTAACAGGCA 1016
DB	100209 CATTATAGGGGAGCTGAATGCAAAAGTAGGAGCAAGAAACACCTGGAGTAACAGGCA 100267


```

Db 102411 TGACTCACAGAAAGACCCCTGATCTGGGAAGATTGAGGTAGGAGAGAGGGGATG 102470
QY 2640 ACAGAGATGAGATGGCTGATGGATCACTGACTCGATGGAGCTGAGTCTGGGTGAATC 2699
Db 102471 ACAGAGATGAGATGGTGGATGGCATCACCGACTCAGTGGACACGGGTTGGGTGGACT 102530
QY 2700 CTGGAGTTGGTATGACACAGGAGGCGCTGTCTCTGGGGCGATTTCATGGGGTTCACAAAGAG 2759
Db 102531 CTGGAGTTGATGATGACAGGAGGCGCTGGCCCTGCTGTGATTCATGGGGTTCACAAAGAG 102590
QY 2760 TTGGACACGACTGAGCAACTGAATGAATGAATGAATGAATGAATGAATGAATGA 2801
Db 102591 TCGGACACGACTGAGCGCTGAATGAATGAATGAATGAATGAATGAATGA 102632

RESULT 34
AC092006
LOCUS Bos taurus clone RP42-456A17, WORKING DRAFT SEQUENCE.
DEFINITION AC092006
ACCESSION AC092006
VERSION 2 GI:21263249
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus

REFERENCE
AUTHORS Akhtar,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C.,
Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J.,
Haghghi,P., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E., Laric,P.,
Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,
Margulies,E.H., Masiello,C., Maskeri,B., Mastrian,S.D.,
McCloskey,J.C., McDowell,J., McGuire,C., Pearson,R.,
Portnoy,M.E., Prasad,A., Schuler,M.G., Stantrispop,S., Thomas,J.W.,
Thomas,P.J., Touchman,J.W., Tsugeon,C., Vogt,J.B., Walker,M.A.,
Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 216698)
Green,E.D.
Direct Submission
Submitted (13-JUN-2001) NIH Intramural Sequencing Center, 8717
Government Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 216698)
Green,E.D.
Direct Submission
Submitted (30-MAY-2002) NIH Intramural Sequencing Center, 8717
Government Circle, Gaithersburg, MD 20877, USA
On May 30, 2002 this sequence version replaced gi:14389328.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoe@ngri.nih.gov
----- Project Information
Center project name: cbj
Center clone name: 456A17

```

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

```

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 216492 bases at least Q40
Consensus quality: 216658 bases at least Q30
Consensus quality: 216689 bases at least Q20
Insert size: 187000; agarose-fp
Quality coverage: 14.77x in Q20 bases; agarose-fp
Quality coverage: 12.74x in Q20 bases; sum-of-contigs

-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 216698: contig of 216698 bp in length.

FEATURES
Location/Qualifiers
1. 216698
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone="RP42-456A17"
/misc_feature 1. 216698
/feature="assembly_fragment"
clone_end:SF6
vector_side:left
clone_end:T7
vector_side:right
1. 42720
/feature="clone overlaps with GenBank Accession Number
AC099745 clone RP42-244N18 (center project name crh)"
120825..216698
/misc_feature 1. 216698
/feature="clone overlaps with GenBank Accession Number
AC092007 clone RP42-73L22 (center project name cbk)"

Query Match 53.5%; Score 1909.81; DB 1; Length 216698;
Best Local Similarity 73.9%; Pred No. 0;
Matches 2473; Conservative 0; Mismatches 257; Indels 618; Gaps 10;

QY 282 AGAGATCTGACAGAATGTGGTCCACTGGAGAAGGAAT-GCAACCACTTCAGTATTCTT 340
Db 17823 AGAGATCTGACAGAATGTGGTCCACTGGAGAAGGAATGGCAAAACCACTTCAGTATTCTT 17852
QY 341 GCCTTGAGAACCCCATGACAGTATGAAAGGCAAAATGATAGGATCTGAAAGAGGAAC 400
Db 17883 GCCTTGAGAACCCCATGACAGTATGAAAGGCAAAATGATAGGATCTGAAAGAGGAAC 17942
QY 401 TCCCAGGTTCAGTAGGTGCCCATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAA 460
Db 17943 TCCCAGGTTCAGTAGGTGCCCATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAA 18002
QY 461 AGAATGAGAGATGAGGCCAAGCAAAAGATATCCAGCTGTGGATCTGACTGGTGATA 520
Db 18003 AGAATGAGAGATGAGGCCAAGCAAAAGATATCCAGCTGTGGATCTGACTGGTGATA 18062
QY 521 TAAGCAAGTCCGATGCTGTAAGAGCAATATTCATAGGAACCTTGAATGTCCAGTCCA 580
Db 18063 GAAGCAAGTCCGATGCTGTAAGAGCAATATTCATAGGAACCTTGAATGTCCAGTCCA 18122
QY 581 TGAATCAAGGCAAAATGGAAGTGGTCAACAAAGAGATGGCAAGAGTGAATGTCAACATTC 640
Db 18123 TGAATCAAGGCAAAATGGAAGTGGTCAACAAAGAGATGGCAAGAGTGAATGTCAACATTC 18182
QY 641 TAGGATCAGCGAAGTAAATGGAATGGGATGAAATTAACCTCAGATGACCAATTATA 700
Db 18183 TAGGATCAGCGAAGTAAATGGAATGGGATGAAATTAACCTCAGATGACCAATTATA 18241

```

QY 701 TCTACTCTGGGCGAGGAATCCCTCAGAGAAATGGAGTAGCCATCATGGTCAACAAAA 760
Db 18242 TCTACTCTGGGCGAGGAATCCCTCAGAGAAATGGAGTAGCCATCATGGTCAACAAAA 18301
QY 761 GAGTCCGAAATGCACTACTGGATGCACTCTCAAAAAACGACAGAAATGATCTCTGTTGTT 820
Db 18302 GAGTCCGAAATGCACTACTGGATGCACTCTCAAAAAATGACAGAAATGATCTCTGTTGTT 18361
QY 821 TCCAGGCAACCAATTCNAATACACAGTAATCCAGTCTATGCCCAACCAAGTAATGCTG 880
Db 18362 TCCAGGCAACCAATTCNAATACACAGTAATCCAGTCTATGCCCAACCAAGTAATGCTG 18421
QY 881 AAGAAGTGTGAAGTTCGAACGGTCTATGAAGACCTTACAAAGACCTTTTAGAACTTAACACCCA 940
Db 18422 AAGAAGTGTGAAGTTCGAACGGTCTATGAAGACCTTACAAAGACCTTTTAGAACTTAACACCCA 18481
QY 941 AAAAAAGATGCTCTCTCAATATATAGGGGACTGGAATGCAAAATGAGGAAGCAAAAGAACAC 1000
Db 18482 AAAAAAGATGCTCTCTCAATATATAGGGGACTGGAATGCAAAATGAGGAAGTCAAGAACAC 18541
QY 1001 CTGAGTAACAGGCAAAATTTGGGCTTCGAATACGGAATGAAGCAGGSCAAAGACTAATAG 1060
Db 18542 TTGAGTAACAGGCAAAATTTGG-CTTGGATACGGAATGAAGCAGGSCAAAGACTAATAG 18600
QY 1061 AGTTTGGCAAGAAATGCACTGCTATAGCAAAACACCTCTTCCAAACACACAGAGAA 1120
Db 18601 AGTTTGGCAAGAAATGCACTGCTATAGCAAAACACCTCTTCCAAACACACAGAGAA 18660
QY 1121 GACTCTACACATGACATCACCAGATGGTCAACACGAAATCAGATTGATTAATCTTT 1180
Db 18661 GACTCTACACATGACATCACCAGATGGTCAACACGAAATCAGATTGATTAATCTTT 18720
QY 1181 GCAGCCAAAGATGAGAGAGTCTATACAGTCAGCAAAACACAGCAGGAGCTTACTGTG 1240
Db 18721 GCAGCCAAAGATGAGAGAGTCTATACAGTCAGCAAAACACAGCAGGAGCTGCTGTG 18780
QY 1241 GCTCAGATCATGAATCTCTTATGCGCAAAATCAGACTTAATTAAGAAAGTAGGGA AAA 1300
Db 18781 GCTCAGATCATGAATCTCTTATGCGCAAAATCAGACTTAATTAAGAAAGTAGGGA AAA 18840
QY 1301 CCAGTACATCACTCAGGTAAAGCTTAATCCATCTCTATGATTAATACAGTGCAGAGTGA 1360
Db 18841 CCAGTACATCACTCAGGTAAAGCTTAATCCATCTCTATGATTAATACAGTGCAGAGTGA 18900
QY 1361 GAAATAGATTTAAGGCTAGATCTGATACAGAGTACTTAATGAACCTATGACAGAGG 1420
Db 18901 GAAATAGATTTAAGGCTAGATCTGATAGAGCGCTGATGAACCTATGATGAGG 18960
QY 1421 TTCATGACATTTACAGGACAGGAGTCCAGACCATCCCATGGAAAGAAATGCAAAA 1480
Db 18961 TTCATGACATTTACAGGACAGGAGTCCAGACCATCCCATGGAAAGAAATGCAAAA 19020
QY 1481 AAGCAAAATGGCTGTCTGGGAGGCTTACAAATAGCTGTGAAAGAGAGAGTGA AAA 1540
Db 19021 AAGCAAAATGGCTGTCTGGGAGGCTTACAAATAGCTGTGAAAGAGAGAGTGA AAA 19080
QY 1541 GCRAAGGAAAGAAAGATTAAGCATCTGATCGAGGTTCCAAA ----- 1588
Db 19081 GCRAAGGAAAGAAAGATTAAGCATCTGATCGAGGTTCCAAAAGATAGCAAGAA 19140
QY 1589 ----- 1588
Db 19141 GCGATAAGAAAGCTTCTTCAGCGATCAATGCAAGAAATAGAGGAAACACAGAGTGG 19200
QY 1589 ----- 1588
Db 19201 GAAAGACTAGAGATCTCTCAAGAAATCAGAGATACCAAGGAAACATTTTCATGCAAGA 19260
QY 1589 ----- 1588
Db 19261 TGGGCTCGATAAGGACAGAAATGGTATGGACCTAACAGNAGCAGAGATATTAGAAGA 19320

QY 1589 ----- 1588
Db 19321 GGTGCAAGAAATACACAGAACTGTACAAAAAGATCTTCATGACCCAGATTAATCAG 19380
QY 1589 ----- 1588
Db 19381 ATGGTGTGATCACTGACCTAGAGCCAGACATCCTGGAATGTGAAGTCAAGTGGGCTTAG 19440
QY 1589 ----- 1588
Db 19441 AAAGCATCACTACGAACAAAGCTAGTGAGGTGATGGAATTCAGCTGAGCTCTTTTCAA 19500
QY 1589 ----- 1588
Db 19501 TCTTAAAGATGACGCTGTGAAGTGTGCTCAATATATGCGCAAAATTTGGAAACATC 19560
QY 1589 ----- 1588
Db 19561 AGCAGTGGCCACAGGACTGGAAGAGGTCAAGTTTCATTCGAATCCAAAGAAAAAGCAATG 19620
QY 1589 ----- 1588
Db 19621 CCAGAAGATGCTCAAACTACCGCACAGTTGCATCTCACATGCTAGTAAAGTAGTGC 19680
QY 1589 -----GAACTTCCAGTTGTTCAAG 1607
Db 19681 TCAAAATTTCTCCAGCCAGGCTTCAGCAATATATGAACCGTGAACCTTCTGATGTTCAA 19740
QY 1608 CTGGTTTATAGAAAGTTCAGAGAACCCAGAGACCAAAATTTGCCAATCTCTGTATCATGG 1667
Db 19741 CTGGTTTATAGAAAGTTCAGAGAACCCAGAGATCGAATTTGCCAAGCTCTGTAGTATCATG 19800
QY 1668 AAAAAAGAGAGTTCAGAAAAACATCTATTTCTGCTTTATGACTATGCAAAAGCCT 1727
Db 19801 AAAAAAGAGAGTTCAGAAAAACATCTATTTCTGCTTTATGACTATGCAAAAGCCT 19860
QY 1728 TTGACTGTGGGGTCACAATAACTGTGGAATTTCTGAAAGGATGGATACAGACC 1787
Db 19861 TTGATGTGTAGATCACAACTCACTGTGGAATTTCTGAAAGAGATGGGATACAGACC 19920
QY 1788 ACCTGACCTGCTTGTAAAAATTTGTATGAGTCAAGGTCAGGAAGCAACAGTTAGAACTGGAC 1847
Db 19921 ACCTGATCTGCTCTTGAGAAATTTGTATGAGTCAAGGTCAGGAAGCAACAGTTAGAACTGGAC 19980
QY 1848 ATGGAACAAACAGACTGGTTCCAGTAGAAGAGATATGTCAGAGCTGTATATTTGTAC 1907
Db 19981 ATGGGACAAACAGACTGGTTCCAAATAGGAAAGAGTACGTCAGAGCTGTATATTTATC 20040
QY 1908 CCGGCTTCTTTAACTTCTATGCAGAG-ACATCATGAGAAACGCTGGGCTGGAAGAACAC 1966
Db 20041 CCGTTTATTTAACTTATGACAGATACATCATGAGAAATGCTGGACTGGAAGAACAC 20100
QY 1967 AAGTGGAAATCAAGATTGCGGGGAGAAATAGCAATAACCTCAGATATGCGAGATGATACCA 2026
Db 20101 AAGTGGAAATCAAGATTGCGGGGAGAAATATCAAAAACCTCAGATATGCGAGATGATACCA 20160
QY 2027 CCGTTATGGCAGAAAGTGAAGAGGACTTAAAGAGCTCTTTGATCAAGGTGAAGAGAGAGA 2086
Db 20161 CCGTTATGAGCAGAAAGTGAAGAGGACTTAAAGAGCTCTTTGATGAAAGTGAAGAGAGAGA 20220
QY 2087 GT-GAAAAAGTGTGCTTAAAGCTCAACATTCAGAAAAACGAGATCATGGCATCTGGTCCC 2145
Db 20221 GTGGAAAGTTGGCTTAAAGCTCAACATTCAGAAAAACGAGATCATGGCATCTGGTCCC 20280
QY 2146 ATCACTTCATCGGAAATAGATGGGAAACAGTGGAAACAGTGCAGACTGTTCAGACTTTTGG 2205
Db 20281 ATCACTTCATCGGAAATAGATGGGAAACAGTGGGAAACAGTGGGAGACTTTA-TTTTGG 20338
QY 2206 GGGGCTCCAAATCACTGCAGATGGTGCAGCCATGAAATTTAAAGACACTTACTCC 2265
Db 20339 GGGGCTCCAAATCACTGCAGATGGTGCAGCCATGAAATTTAAAGAGTGCITACTCC 20398
QY 2266 TTGAAGAAAAAGTTA--ACCNAACCTAGATGATATTGAAAAAGCAGAGACATTAACCTTG 2323

Db 20399 TTGGAAGAAATTTATGACCAACCTAGATAGCATATTGAAAGCAGACATTAATTGTC 20458
Qy 2324 CAACAAGCCCATCTAGTCAAGCTATGTTTTCCAGTGGTCATGTGATGTCAGA 2383
Db 20459 CAACAAGGTCATCTAGTCAAGCTATGTTTTCTCTGGTTCATGTGATGTCAGA 20518
Qy 2384 GTTGACTGTGAAGAAAGCTGAGCAGTCAAGAAATTTGATGTTTGAACCTGTGTGTTGGA 2443
Db 20519 GTTGACTGTGAAGAAAGCTGAGTGTCTGAAGAAATTTGATGTTTGAACCTGTGTGTTGGA 20578
Qy 2444 GAAGACTTTGAGAGTCCCTTGGACTGCAAGGAGATCCACCACTGATCCATCTCAAGAGAGA 2503
Db 20579 GAAGACTTTGAGAGTCCCTTGGACTGCAAGGAGATCCACCACTGATCCATCTCAAGAGAGA 20638
Qy 2504 TCAGCCCTGGGATTTCTTTTGAAGGAATGATGCTTAAAGCTGAACTCCAGTACTTTGGCC 2563
Db 20639 TCAGCCCTGGGATTTCTTTTGAAGGAATGATGCTTAAAGCTGAACTCCAGTACTTTGGCC 20698
Qy 2564 ACCTGATCAGAGAGTCACTCACTGGAAGAACCCCTGATGCTGGAGGAGATTTGGGGGCA 2623
Db 20699 ACCTCATGCGAAGAGTTGATCTCATTTGGAAAGACTCTGATGCTGGAGGAGATTTGGGGGCA 20758
Qy 2624 GGAGGAGAGGGGACGACAGAGGATGATGGCTGGATGCTGATGCATCACTGACTCGATGGACG 2683
Db 20759 AGAGGAGAGCGGACGACAGAGATGATGGCTGGATGCTGATGCATCACTGACTCGATGGACG 20818
Qy 2684 TGAGTCTGGGTGAACCTCTCGAGTGTGTGATGACAGGAGGCTGTCTCTGGCGGATTC 2743
Db 20819 TGAGTCTGAGTGAACCTCTCGAGTGTGTGATGACAGGAGGCTGTCTCTGGCGGATTC 20878
Qy 2744 ATGGGGTCANAGAGTGGACAGCTGACGACTGACCACTGAACTGAACTGACTGACTGAAA 2803
Db 20879 ACGGGGTCGAAAGAGTCGACAGCTGACGAGCTGACGAGCTGATCTGATCTGATCTG 20930
Qy 2804 CCTAGTAGTTTATTTACTCAGAAAATAGTAAATTTTCAATGATGATTTCAAAATTTATTCAT 2863
Db 20931 -----AAGTTTATTACACGTTGTTAGTTTCAGTGTACATGATGATTCAGT 20983
Qy 2864 AATGTTGGTTAAGATAATAGATTTTC-AAATGATTTTATCTTTGATTTTCTTACT 2922
Db 20984 CATACATATACATACATTCATTTCTTCAAAATTTCTTCTATATAGGTTTTCACAGAA 21043
Qy 2923 TATTTAATTTGGGATTTTAACTATTTCTCAATGACTTCTTATTTCTAATATTTACTTAT 2982
Db 21044 TATTAATATAGACCTGATATTCAGTAGGTTCTGTTCTCAATATCTGCTTACATAT 21103
Qy 2983 TCTATTTACTTTAATGCACTTATTTTATTTATTTTCTAATAAAA 3030
Db 21104 AGTAGTGTATATCTAGGAAATGTTTTTATTGTTTGGCTGTTATA 21151

RESULT 35
AC099746 Bos taurus clone RP42-244N18, WORKING DRAFT SEQUENCE, 4 ordered
LOCUS AC099746 171875 bp DNA linear HTG 30-MAY-2002
DEFINITION Bos taurus clone RP42-244N18, WORKING DRAFT SEQUENCE, 4 ordered
pieces.
AC099746
VERSION AC099746.2 GI:21263247
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos
1 (bases 1 to 171875)
REFERENCE Akhter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C.,
Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J.,
Haghghi, P., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Laric, P.,
Lee-Lin, S.-O., Legaspi, R., Maduro, O.L., Maduro, V.B.,
Margulies, E.H., Masello, C., Maskeri, B., Mastrian, S.D.,
McCloskey, J.C., McDowell, J., McGuire, J., Pearson, R.,

Portnoy, M.E., Prasad, A., Schueler, M.G., Stantripp, S., Thomas, J.W.,
Thomas, P.J., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A.,
Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 171875)
Green, E.D.
Direct Submission
Submitted (20-NOV-2001) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 171875)
Green, E.D.
Direct Submission
Submitted (30-MAY-2002) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
On May 30, 2002 this sequence version replaced gi:17017550.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@nih.gov
----- Project Information
Center project name: crh
Center clone name: 244N18

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 171182 bases at least Q40
Consensus quality: 171457 bases at least Q30
Consensus quality: 171549 bases at least Q20
Insert size: 146000; agarose-fp
Insert size: 171575; sum-of-contigs
Quality coverage: 12.65x in Q20 bases; agarose-fp
Quality coverage: 10.76x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced
* By the finished sequence as soon as it is available and the accession number will be preserved.

1 123482: contig of 123482 bp in length
123582: gap of unknown length
137343 137342: contig of 13760 bp in length
137443 137442: gap of unknown length
142843 142844: contig of 5402 bp in length
142845 142844: gap of unknown length
142945 171875: contig of 28931 bp in length.

Location/Qualifiers
1. 171875
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone_RP42-244N18"
/clone_lib="RP42"
1. 123482
misc_feature

/note="assembly_fragment clone_end:r7 vector_side:left 1. 22552 /note="clone overlaps with GenBank Accession Number AC099407 clone RP42-270J1 (center project name crg)" 123583. 137342 /note="assembly_fragment" 129328. 171875 /note="clone overlaps with GenBank Accession Number AC092006 clone RP42-456A17 (center project name cbj)" 137443. 142844 /note="assembly_fragment" 142945. 171875 /note="assembly_fragment clone_end:SP6 vector_side:right"		Query Match 53.4%; Score 1907.01; DB 1; Length 171875; Best Local Similarity 73.0%; Pred. No. 0; Matches 2499; Conservative 0; Mismatches 300; Indels 624; Gaps 10;	
misc_feature	282	AGAGATCTGACAGAAATGGTCCACTGGAGAGGGAATGCGAATCTGAAAGAGGAAC	400
misc_feature	146977	AGAGATCTGACAGAAATGGTCCACTGGAGAGGGAATGCGAATCTGAAAGAGGAAC	147036
misc_feature	341	GCCTTGAGAACCCATGAACAGTATGAAAGGCCAAATGATAGGATCTGAAAGAGGAAC	400
misc_feature	147037	GCCTTGAGAACCCATGATAGTACGAAAGGCCAAATGATAGGATCTGAAAGAGGAAC	147096
misc_feature	401	TCCCCAGTCTAGTGGTCCCATATGCTACTGGAGATCTGAGTGGAGAAATCACTCAGAA	460
misc_feature	147097	TCCCCAGTCTAGTGGTCCCATATGCTACTGGAGATCTGAGTGGAGAAATCACTCAGAA	147156
misc_feature	461	AGAATGAGAGATGGCCCAAGCAAAAGAAATACCCAGCTGTGGATGTCAGCTGGTGATA	520
misc_feature	147157	AGAATGAGAGATGGCCCAAGCAAAAGAAATACCCAGCTGTGGATGTCAGCTGGTGATA	147216
misc_feature	521	TAAGCAAGTCCGATGCTGTAAGAGCAATATTCATAGGAACCTGGAATGTCAGGTCCA	580
misc_feature	147217	GAAGCAAGTCCGATGCTGTAAGAGCAATATTCATAGGAACCTGGAATGTCAGGTCCA	147276
misc_feature	591	TGAATCAAGGCAATGGAATGTCATAAAGAGATGCAAGAGTGGCAAGTGAATGTCACATTC	640
misc_feature	147277	TGAATCAAGGCAATGGAATGTCATAAAGAGATGCAAGAGTGGCAAGTGAATGTCACATTC	147336
misc_feature	641	TAGGAATCAGGCAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT	700
misc_feature	147337	TAGGAATCAGGCAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT	147395
misc_feature	701	TCTACTACTGGGGCAGGAATCCCTCAGAGAAATGGAGTAGCCATCATGGTCAACAAAA	760
misc_feature	147396	TCTACTACTGGGGCAGGAATCCCTCAGAGAAATGGAGTAGCCATCATGGTCAACAAAA	147455
misc_feature	761	GAGTCCGAATGCACTCTCGATGCACTCTCAAAAGCAGAGATGATCTCTGTTGTT	820
misc_feature	147456	GAGTCCGAATGCACTCTCGATGCACTCTCAAAAGCAGAGATGATCTCTGTTGTT	147515
misc_feature	821	TCCAGGCAAAACCAATTCATATCAGTAATCCAAAGTCTATGCCCCAACCAAGTAATGCTG	880
misc_feature	147516	TCCAGGCAAAACCAATTCATATCAGTAATCCAAAGTCTATGCCCCAACCAAGTAATGCTG	147575
misc_feature	881	AAGAAGCTGAAGTTGAACGGTCTTATGAGACCTACAGACCTTTTAGAATCAACCCCA	940
misc_feature	147576	AAGAAGCTGAAGTTGAAGTCTTATATAAAGACCTACAGACCTTTTAGAATCAACCCCA	147635
misc_feature	941	AAAAAGATGCTCTTCTATTATAGGGACTGGAATGCAAAAGTAGGAAGCAAGAAACAC	1000
misc_feature	147636	AAAAAGATGCTCTTCTATTATAGGGACTGGAATGCAAAAGTAGGAAGTCAAGAAACAC	147695
misc_feature	1001	CTGGAGTAAACAGGCAAAATTTGSCCTTGGAAATACGGAATGAGCAGGCGCAAGACTATAG	1060
misc_feature	147696	TTGGAGTAAACAGGCAAAATTTGSCATGGGAATACGGAATGAGCAGGCGCAAGACTATAG	147755

QY	1061	AGTTTTGCCAAGAAAATGCACTGGTCTAGCAAAACACCCCTCTTCCAAACACACAGAGAA	1120
Db	147756	AGTTTTGCCAAGAAAATGCACTGGTCTAGCAAAACACCCCTCTTCCAAACACACAGAGAA	147815
QY	1121	GACTCTACACATGGACATCACAGATGGTCAACACCGAAATCAGATTATATCTTT	1180
Db	147816	GACTCTACACATGGACATCACAGATGGTCAACACCGAAATCAGATTATATCTTT	147875
QY	1181	CGACCAAGATGGAGAGCTCTATACAGTCAAGCAAAACCAAGACCGAGCTTACTGTG	1240
Db	147876	CGACCAAGATGGAGAGCTCTATACAGTCAAGCAAAACCAAGACCGAGCTTACTGTG	147935
QY	1241	GCTCAGATCATGAACCTCTTATGCAAAATTCAGACTTAAATTTGAAGAAATGAGGAAAA	1300
Db	147936	GCTCAGATCATGAACCTCTTATGCAAAATTCAGACTTAAATTTGAAGAAATGAGGAAAA	147995
QY	1301	CCACTAGATCACTCAGGTAAGACCTAAATCCATCCCTTATGATTATACCTGGAAGTGA	1360
Db	147996	CCACTAGATCACTCAGGTAAGACCTAAATCCATCCCTTATGATTATACCTGGAAGTGA	148055
QY	1361	GAATAGATTTAAGGCGCTAGATCTGTATGACAGAGTACCTAATGAACATATGGACAGG	1420
Db	148056	GAATAGATTTAAGGCGCTAGATCTGTATGACAGAGTACCTAATGAACATATGGACAGG	148115
QY	1421	TTTCATGACATTTGACAGAGACAGGATCGAGACCATCCCATGGAAAGAAATGCAAAA	1480
Db	148116	TTTCATGACATTTGACAGAGACAGGATCGAGACCATCCCATGGAAAGAAATGCAAAA	148175
QY	1481	AGCAAAATGGTGTCTGGGAGGCTTTACAAATAGCTGTGAAAAGAGAGAGAGTGAATA	1540
Db	148176	AGCAAAATGGTGTCTGGGAGGCTTTACAAATAGCTGTGAAAAGAGAGAGAGTGAATA	148235
QY	1541	GCAAGGAAAAAGGAAAGATAAAAGCATGCAATGCAAGTTCCTCAAA	1588
Db	148236	GCAAGGAAAAAGGAAAGATAAAAGCATGCAATGCAAGTTCCTCAAAAGAAATGCAAGAA	148295
QY	1589	-----	1588
Db	148296	GCAATAGAAAGCCCTTCTCAGCGATCAATGCAAGAAATGAGGAAACACAGATGG	148355
QY	1589	-----	1588
Db	148356	GAAAGACTAGAGATCTCTTCAAGAAATCAGAGATACCAAGGAACATTTTCATGCAAGA	148415
QY	1589	-----	1588
Db	148416	TGGGCTCGATAAAGCAGAGAAATGGTATGGACCTAACAGAGAGAGAGATATTAAGAAGA	148475
QY	1589	-----	1588
Db	148476	GGTGGCAAGATACACAGAGAACTGTACAGAAAGATCTTCACGCCAGATATCAAG	148535
QY	1589	-----	1588
Db	148536	ATGGTGTGATCACTGATAGAGCCAGACATCCTGGAATGTGAAGTCAAGTGGGCTTAG	148595
QY	1589	-----	1588
Db	148596	AAAGCATCACTAGAACAAAGCTAGTGGAGTGGAAATTCAGCTTTCAGCTTTCATAA	148655
QY	1589	-----	1588
Db	148656	TCTAAAGAGATGATGCTGTGAAAGTGTGCACTCAATATGCCAGCAAAATTTGGAACCTC	148715
QY	1589	-----	1588
Db	148716	AGCAGTGGCCACAGGACTGGAAAGGTCAGTTTTCATTTCCAAATCCCAAGAAACCAATG	148775
QY	1589	-----	1588
Db	148776	CCAAAGATGCTCAAACTACCAACACAGATTGCACTCATCTCATCTAGTAAAGTAGTGC	148835

----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: <http://www.nisc.nih.gov>
 Contact: nisc_zoon@nigri.nih.gov
 ----- Project Information
 Center project name: djx
 Center clone name: 246J24

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 182395 bases at least Q40
 Consensus quality: 182646 bases at least Q30
 Consensus quality: 182726 bases at least Q20
 Insert size: 140000; agarose-fp
 Insert size: 182748; sum-of-contigs
 Quality coverage: 10.60x in Q20 bases; agarose-fp
 Quality coverage: 8.12x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 1 49758: contig of 49758 bp in length
 49759 49858: gap of unknown length
 49859 120040: contig of 70182 bp in length
 120041 120140: gap of unknown length
 120141 169454: contig of 49314 bp in length
 169455 169554: gap of unknown length
 169555 183048: contig of 13494 bp in length.

FEATURES

Source

1..183048
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9913"
 /clone="RP42-246J24"
 /clone_lib="RP42"

misc_feature

1..49758
 /note="assembly_fragment"
 clone_end:SP6
 vector_side:left
 49859..120040
 /note="assembly_fragment"
 120141..169454
 /note="assembly_fragment"
 169555..183048
 /note="assembly_fragment"
 clone_end:T7
 vector_side:right

misc_feature

49859..120040
 /note="assembly_fragment"

misc_feature

120141..169454
 /note="assembly_fragment"

misc_feature

169555..183048
 /note="assembly_fragment"

Query Match

Best Local Similarity 52.9%; Score 1891.1; DB 1; Length 183048;
 Matches 2361; Conservative 0; Mismatches 164; Indels 603; Gaps 7;

QY 272 GTCTGTTGGAGAGATCTGACAGAAATGTGGTCCACTGGAGAGGGAAT-GCAAACCACTT 330
 Db 41609 GTTTAGATAGACAGTTTGGACAGAAATGTGGTCCACTGGAGAGGGAATGGCAAACCACTT 41668
 QY 331 CAGTATTTCTTGGCTTGAAGAAACCCATGACAGTATGAAGGCAAAATGATAGATAGTCTG 390
 Db 41669 AAGTATTTCTTGGCTTGAAGAAACCCATGACAGTATGAAGGCAAAATGATAGATAGTCTG 41728
 QY 391 AAAGAGGAATCCCCAGGTCAAGTAGTCCCATATGCTACTGGAGATCACTGGAGAAAT 450
 Db 41729 AAAGAGGAATCCCCAGGTCAAGTAGTCCCATATGCTACTGGAGATCACTGGAGAAAT 41788
 QY 451 AACTCCAGAAAGAAATGAAGAGATGAGGCCAAAGCAAAAGAAATACCCAGCTGTGGATGTG 510
 Db 41789 AACTCCAGAAAGAAATGAAGAGATGAGGCCAAAGCAAAAGAAATACCCAGCTGTGGATGTG 41848
 QY 511 ACTGGTGATATAAGCAAGGTCCGATGCTGTAAAGAGCAATATTGTCATAGGAACCTGGAAT 570
 Db 41849 ACTGGTGATATAAGCAAGGTCCGATGCTGTGTAAAGAGCAATATTGTCATAGGAACCTGGAAT 41908
 QY 571 GTCAAGTCCATGAATCAAGGCAAAATTTGGAAGTGTGTCAAACCAAGAGATGGCAAGAGTGAAT 630
 Db 41909 GTCAAGTCCATGAATCAAGGCAAAATTTGGAAGTGTGTCAAACCAAGAGATGGCAAGAGTGAAT 41968
 QY 631 GTCAACATTTAGGATCAGGCAAAATTAATGGAATGGATGGATGGATGGATGGATGGATGGAT 690
 Db 41969 GTCAACATTTAGGATCAGGCAAAATTAATGGAATGGATGGATGGATGGATGGATGGATGGAT 42028
 QY 691 GACCAATTATCTACTACTGCGGCGAGGAATCCCTCAGAAGAAATGGAGTAGCCATCATG 750
 Db 42029 GAGCAATTATTTACTACTGCGGCGAGGAATCCCTCAGAAGAAATGGAGTAGCCATCATG 42088
 QY 751 GTCAACAAAGAGTCCGAAATTCAGTACTTGGATGCGAGTCTCAAAACGACAGATGATC 810
 Db 42089 GTCAACAAAGAGTCCGAAATTCAGTACTTGGATGCGAGTCTCAAAACGACAGATGATC 42148
 QY 811 TCTGTTTCTTCCAAAGGCAAAACCAATCAATATCAAGTAATCCAAAGTCTATGCCCCAACCC 870
 Db 42149 TCTGTTTCTTCCAAAGGCAAAACCAATCAATATCAAGTAATCCAAAGTCTATGCCCCAACCC 42208
 QY 871 -AGTAATGCTGAAGAGTGAAGTGAACGGTCCCTATGAAGACCTACAGACCTTTTAGA 929
 Db 42209 AAGTAACGCTGAAGAGTGAAGTGAACGGTCCCTATGAAGACCTACAGACCTTTTAGA 42268
 QY 930 ACTACACCCCAAAAGAGTGTCTTCTCATTATAGGGGACCTGGAATCCAAAGTAGGAAG 989
 Db 42269 ACTACACCCCAAAAGAGTGTCTTCTCATTATAGGGGACCTGGAATCCAAAGTAGGAAG 42328
 QY 990 CAAAGAAACACCTGGAGTAACAGGCAAAATTTGGCTTTGGAAATACGGAATGAAGCAGGCA 1049
 Db 42329 TCAAGAAACACCTGGAGTAACAGGCAAAATTTGGCTTTGGAAATACGGAATGAAGCAGGCA 42388
 QY 1050 AAGACTAATAGAGTTTCCCAAGAAATGCACTGGTCTATAGCAACCAACCTCTTCCACACA 1109
 Db 42389 AAGACTAATAGAGTTTCCCAAGAAATGCACTGGTCTATAGCAACCAACCTCTTCCACACA 42448
 QY 1110 ACACAAGAGAGAGTCTTACATGACATCACAGATGGTCAACACCCGAAATCAGATTGA 1169
 Db 42449 ACACAAGAGAGAGTCTTACATGACATCACAGATGGTCAACACCCGAAATCAGATTGA 42508
 QY 1170 TTATATTCTTTGGAGCCAAAGATGGAGAGTCTTATACAGTACAGCAAAACCAAGCAGG 1229
 Db 42509 TTATATTCTTTGGAGCCAAAGATGGAGAGTCTTATACAGTACAGCAAAACCAAGCAGG 42568
 QY 1230 AGCTTACTGTGGCTCAGATCATGAATCCCTTATTGGCAAAATTCAGATTTAAATGAAGAA 1289
 Db 42569 AGCTTACTGTGGCTCAGATCATGAATCCCTTATTGGCAAAATTCAGATTTAAATGAAGAA 42628
 QY 1290 AGTAGGAAAAACCACTAGATCACTCAGGTAAAGACCTAAATCCCAATCCCTTTATGATTATAC 1349
 Db 42629 AGTAGGAAAAACCACTAGATCACTCAGGTAAAGACCTAAATCCCAATCCCTTTATGATTATAC 42688
 QY 1350 AGTGGAACTGAGAAATAGATTTAAGGGCCTAGATCTGTATAGACAGAGTACCTAATGAAT 1409

42689 ATGGAGTGAAGATAGATTGAAGGCTTAGATCTGATAGATAGAGTGCCTGATGACT 42748
1410 ATGGACAGAGTTTCATGACATTTGTACAGGAGACAGGATCGAGACCATCCCATGGAAAA 1469
42749 ATGGAATGAGTTCGTGAAGTTGTACAGGAGACAGGAATCAAGACCATCCCATAGAAAA 42808
1470 GAAATGCAGAAAGCAAAATGGCTGTCTGGGAGGCTTACAAATAGCTGTGAAAGAG 1529
42809 GAAATGAAAGAAAGCAAAATGGCTGTCTGGGAGGCTTACAAATAGCTGTGAAAGAG 42868
1530 AGAAGTGAAGCAAGCAAGAAAGAAAGAAAGATGAAAGCATCTGAAATGCAGAGTTCCAAA- 1588
42869 AGAAGTGAAGCAAGCAAGAAAGAAAGAAAGATGAAAGCATCTGAAATGCAGAGTTCCAAAG 42928
1589 ----- 1588
42929 AATAGCAAGGAGATAGAAAGCTTCTTTCAGCAATCAATGCAAGAAATAGAGGAAAA 42988
1589 ----- 1588
42989 CAACAGATGGAAAGACTAGGATCTCTTCAAGAAATCAGAGATACCAAGAACATTT 43048
1589 ----- 1588
43049 TCCTGCAAGATGAGCTCGATAGAAAGACAGAAATGGTATGCACTAACAGAGCAGAAGA 43108
1589 ----- 1588
43109 TATTAGAGAGATGGCAAGAAATACACAGAAAGAACTATACAAAAGATCTTCATGACCC 43168
1589 ----- 1588
43169 AGATAATCAGTCTGTGATCACTGACCTAGACCAGACATCCTGGAAATGTGAAGTCAA 43228
1589 ----- 1588
43229 GTGGGCTTAGAAAGCATCACTAGAAACAAAGCTAGTGGAGGCGATGGAATCCAGTTGA 43288
1589 ----- 1588
43289 GCTATTCCAAATCCTGAAAGATGATGCTGTGAAAGTCTGCACCTCAATATGCCAGCAAT 43348
1589 ----- 1588
43349 TTGGAAGCACTCAGCAGTGGCCACAGGACTGGAAAGGTGAGTTTTCATTCATCCAAATCCCAA 43408
1589 ----- 1588
43409 GAAAGGCAATGCCAAAGATGCTCAAACTACCACAAATGGCACTCATCTCACATGCTAG 43468
1589 ----- 1596
43469 TAAAGTAAAGCTCAAAATTTCTCCAAAGCCAGCTTCAGCAATATGTGAAACCGTGAACCTCC 43528
1597 AGTTGTTCAAGCTGTTTAGAAAGTTCAGAGAACAGAGACCAATTCGCAACATCCT 1656
43529 AGATATTCAAGCTGTTTAGAAAGGAGAGAACAGAGATCAAAATGTGCAACATCTG 43588
1657 CTGTATCATGGAAGAAAGCAAGAGAGTTCAGAAAGAAATCTATTCTGCTTTATGACTA 1716
43589 CTGGATCATTTGAAAGCAAGAGAGTTCAGAGAAATCTATTCTGTTTATGACTA 43648
1717 TGCAAAAGCCTTTGACTGTGGGGTCAATTAATACTGTGAAATTTCTGAAAGGATGGG 1776
43649 TGCCAAAGCCTTTGACTATGTGATCAAAATTAATACTGTGAAATTTCTGAAAGGATGGG 43708
1777 AATACCAGACCACTGACCTGACTCTTTGAAATTTTGTATGAGGTTCAGGAGCAACAGT 1836
43709 AATACCAGACCACTGACCTGCTCTTTGAGAAATTTGTATGAGGTTCAGGAGCAACAT 43768
1837 TAGAACTGGACATGGAACACAGACTGTTTCCAAAGTAGGAAAGGAGTATGTCAAGGCTG 1896

43769 TAGAACTGGACAGG---AACAGACTGGTTCCAAATAGGAAAGGAGTTCGTCAAGGCTG 43825
1897 TATATTGTACCCGCTGTTTAACTTCTATGACAG--ACATCATGAGAAACGCTGGGCT 1955
43826 TATATTGTACCTGTTTATTATTAACATATATGACAGTACATCATGAGAAACGCTGGACT 43885
1956 GGAAGAGCACAAAGCTGGAATCAAGATTGCCGGGAGAAATAGCAATAACCTCAGATATGC 2015
43886 GGAAGAAACACAACTGGAATCAGATTGCCGGGAGAAATATCAATACTCAGACATGC 43945
2016 AGATGATACCCCTTTATGGCAAGAAAGTGAAGAGAACTAAAAAGCCTCTTTGATGAAGT 2075
43946 AGATGACACCCCTTTATGGCAAGAAAGTGAAGAGAAATTCAAAAGCCTCTTTGGTGAAGT 44005
2076 GAAAGAGGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTTCAGAAACGAGATCATGC 2135
44006 GAAAGTGAAGTGAAGAAAGTTGGCTTAAAGCTCAACATTTCAGAAACGAGATCATGC 44065
2136 ATCTGCTCCCACTCTCATGGAAATAGATGGGAAAACAGTGGAAACAGTGTCCAGACTT 2195
44066 ATCCGCTCCCACTCTCATGGAAATAGATGGGAAAACAGTGGAAACAGGCTCAGACTT 44125
2196 TATTTTGGGGGCTCCAAATCACTGCAGATGCTGAGTGCAGCATGAATTTAAAGA 2255
44126 TATTTT--CTGGGCTCCAAATCACTAGATGCTGAGTGCAGCATGAATTTAAAGA 44183
2256 CACTTACTCTTGGAAAGAAAGTTA--ACCACCTTAGATAGCATATTGAAAAGCAGAGAC 2313
44184 CACTTACTCTTGGAAAGAAAGTTATGACCACTTAGATAGCATATTCAAAGCAGAGAC 44243
2314 ATTACTTGCCTTGGAAAGCCTCATCTAGTCAAGGCTATGTTTTCAGTGTGATGAT 2373
44244 ATTACTTGCCTTGGAAAGCCTCATCTAGTCAAGGCTATGTTTTCCTGTGCTATGAT 44303
2374 GGATGTGAGTGGACTGTGAAGAAAGCTGAGCACTGAAGAAATGATGCTTTTCAACTG 2433
44304 GGATGTGAGTGGACTGTGAAGAAAGCTGAGTCCGAGAAATGATGCTTTTCAACTG 44363
2434 TGGTGTGAGAGAACTCTTGAGAGTCCCTTGGAAGTCCGAGAGATCCAAACAGTCCATT 2493
44364 TGGTGTGAGAGAACTCTTGAGAGTCCCTTGGAAGTCCGAGAGATCCAAACCAATCCATT 44423
2494 CTGAAGGAGATCAGCCCTGGGATTTCTTGGAGGAATGATGCTAAAGTGAATCCAG 2553
44424 GTGAAGGAGATCAGCCCTGGGATTTCTTGGAGGAATGATGCTAAAGTGAATCCAG 44483
2554 TACTTTGGCCACTGATCAGAAAGAGTCACTCACTGGAAAAGACCCCTGATGCTGGAGGG 2613
44484 TACTTTGGCCACTCATGCGAAGTGTGATTCTTTGGGAAGAGCTCTGATCTGGAGGG 44543
2614 ATTGGGGCAGAGAGAGAGGAGCAGCAGAGGATGAGATGCTGGATGGCATCACTGAC 2673
44544 GTTGGGGCAGAGAGTGAAGGGGACGACAGAGGATGAGATGCTGGATGGCAACACTGAC 44603
2674 TCGATGAGCTGAGTCTGGGTGAACTCCTGAGATTTGGTGTGATGACAGAGGAGGCTGTCT 2733
44604 TCGATAGAGCTGAGTCTGGGTGAACTCCTGAGATTTGGTGTGATGACAGAGGAGGCTTGGCT 44663
2734 GCGGGGATTCATGGGGTCAAAAGAGTGGACACGACTGAGCAACTGAACTGAAC 2793
44664 GCTGGGATTCATGGGGTTCGAAAGAGTGGACACGACTGAGCAACTGATCTGATCTGATC 44723
2794 TGTACTGA 2801
44724 TGATCTGA 44731

RESULT 37

AC108896/c

LOCUS

DEFINITION

AC108896

ACCESSION

223717 bp DNA linear HTG 20-JUL-2002
Bos taurus clone RP42-89G3, WORKING DRAFT SEQUENCE, 5 ordered
pieces.


```

Db      89415 AAGBAGAACTAAGAGCCTTTGATGAAAGGAAAGGAAAGTAAAGAGTTGGCTGAA 89356
Qy      2105 AGCTCAACATTCAGAAACGAGATCATGGCATCTGTGCCATCACATTCATGGAATAG 2164
Db      89355 AGCTCAATATTCAGAAATGAAGATCATGGCATCTGTGCCATCACATTCATGGAATAG 89296
Qy      2165 ATGGGGAACAGTGGAAACAGTGCAGACATTTATTTTGGGGGCTCCAAATCACTGC 2224
Db      89295 ATGGGGAATAGTCGAACAGTGGCTGACATTTATTTCT--TGGGCTCCAAATCACTGC 89238
Qy      2225 AGATGGTGAAGTGCAGCCATGAAATTAAGACACTTACTCTCTGGAGAGAAAGTTA--AC 2282
Db      89237 AGATGGTGAAGTGCAGCCATGAAATTAAGACACTTACTCTCTGGAGAGAAAGTTATGAT 89178
Qy      2283 CAACCTAGATAGCATATTGAAGACAGAGACATTAACCTTCCACAAAGCCCATCTAGT 2342
Db      89177 CAACCTAGATAGCATATTGAAGACAGAGACATTAACCTTCCACAAAGGTCCTAGT 89118
Qy      2343 CAAGCTATGCTTTTTCAGTGGTCACTGATGATGATGGATGGAAGTGGGACTGTGAAGAAAGC 2402
Db      89117 CAAGCTATGCTTTTTCAGTGGTCACTGATGATGATGGATGGAAGTGGGACTGTGAAGAAAGC 89058
Qy      2403 TGAGCACTGAAGATGATGCTTTTGAACCTGTGTTTGAAGAGACTCTTGAGAGTCCC 2462
Db      89057 TGAGCACTGAAGATGATGCTTTTGAACCTGTGTTTGAAGAGACTCTTGAGAGTCCC 88998
Qy      2463 TTGACTGCAAGGAGATCCACCAAGTCCATTCGAAGAGATCAGCCCTGGGATTTCTTT 2522
Db      88997 ATGACTACAAGGAGATCCACCAAGTCCATTCGAAGAGATCAGCCCTGGGATTTCTTT 88938
Qy      2523 GGAAGGAATGATGCTAAAGTGAACCTCCAGTACTTTGGCCACTGATCAGAAGAGCTGA 2582
Db      88937 GGAAGGAATGATGCTAAAGTGAACCTCCAGTACTTTGGCCACTGATCAGAAGAGCTGA 88878
Qy      2583 CTCAGTGAAGAACCCCTGATGCTGGGAGGATTTGGGCGAGGAGAGAGGAGGACACA 2642
Db      88877 CTCAGTGAAGAACCCCTGATGCTGGGAGGATTTGGGCGAGGAGAGAGGAGGACACA 88818
Qy      2643 GAGGATGAGTGGGTGGATGGCATCACTGATCGATGAGCTGAGTCTGGGTGAACTCCCT 2702
Db      88817 GAGGATGAGTGGGTGGATGGCATCACTGATCGATGAGCTGAGTCTGGGTGAACTCCCT 88758
Qy      2703 GGAGTTGATGATGACAGAGGAGGCTGCTCGGGCGATTCATGGGCTCAAGAGTTG 2762
Db      88757 GGAGTTGATGATGACAGAGGAGGCTGCTCGGGCGATTCATGGGCTCAAGAGTTG 88699
Qy      2763 GACAGCTGAGCAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTTATATAC 2822
Db      88698 GACATGACTGAGCAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTTATATAC 88639
Qy      2823 TCAGAAATAGTAATTTCA 2843
Db      88638 TCTGGAGAATTAATGACAGA 88618

RESULT 28
AC108176/c
LOCUS   AC108176 154440 bp DNA linear HTG 20-JUL-2002
DEFINITION Bos taurus clone RP42-57H8, WORKING DRAFT SEQUENCE, 2 ordered
          pieces.
ACCESSION AC108176
VERSION   AC108176.2 GI:21913014
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE   Bos taurus (cow)
ORGANISM Bos taurus
REFERENCE
AUTHORS  Akhter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,
          Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C.,
          Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J.,
          Hadhighi, P., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Lari, P.,
          Lee-Lin, S.-Q., Legaaspi, R., Maduro, Q.L., Maduro, V.B.,
          Marculescu, E.H., Mastello, C., Maskeri, B., Mastrian, S.D.,
          McCloskey, J.C., McDowell, J., Faguirgan, C., Pearson, R.,
          Portnoy, M.B., Prasad, A., Schueler, M.G., Stantrop, S., Thomas, J.W.,
          Thomas, P.J., Touchman, J.W., Turgeson, C., Vogt, J.L., Walker, M.A.,
          Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.
          NISC Comparative Sequencing Initiative
          Unpublished
          2 (bases 1 to 154440)
          Green, E.D.
          Direct Submission
          Submitted (26-JAN-2002) NIH Intramural Sequencing Center, 8717
          Grovemont Circle, Gaithersburg, MD 20877, USA
          3 (bases 1 to 154440)
          Green, E.D.
          Direct Submission
          Submitted (20-JUL-2002) NIH Intramural Sequencing Center, 8717
          Grovemont Circle, Gaithersburg, MD 20877, USA
          On Jul 20, 2002 this sequence version replaced gi:18376875.
          ----- Genome Center
          Center: NIH Intramural Sequencing Center
          Center code: NISC
          Web site: http://www.nisc.nih.gov
          Contact: nisc_zoo@nhgri.nih.gov
          ----- Project Information
          ----- Project name: axg
          Center clone name: 057H08

The sequence data in this record represents an 'enhanced'
version of a Phase 2 submission. Specifically, the indicated
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8X average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associated
with a Phrap-derived quality score.
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 154258 bases at least Q40
Consensus quality: 154312 bases at least Q30
Consensus quality: 154337 bases at least Q20
Insert size: 146000; agarose-fp
Insert size: 154340; sum-of-contigs
Quality coverage: 13.05x in Q20 bases; agarose-fp
Quality coverage: 12.34x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 96915: contig of 96915 bp in length
* 96916 97015: gap of unknown length
* 97016 154440: contig of 57425 bp in length.
          Location/Qualifiers
            1..154440
              /organism="Bos taurus"
              /mol_type="genomic DNA"
              /db_xref="taxon:9913"
              /clone="RP42-57H8"
              /clone_lib="RP42"
              1..96915
              /note="assembly_fragment"
          misc_feature

```

clone_end:T7 vector_side:left		840	TATCAGAGTAATCCAAAGTCTATGCCCCAAACCAAGTAACTGCTGAAGAGCTGAAGTTGAACG	899	
1..64027		134118	TATCAGTCTAATCCAAAGTCTATGCCCCAAACCAAGTAACTG-AGAAGCTTAAATTTGAACA	134060	
/notes="clone overlaps with GenBank Accession Number AC108175 clone RP42-250120 (center project name axf)"		900	GTCCCTATGAGACCTACAGACCTTTTAGAACTAACACCCCAAAAGAGTGTCTCTCTCAT	959	
42829..154440		134059	GTCTTATGAAGACCTTATAAGACCTTCTAGAACTCACACCCCAAAAGAGTGTCTCTCTCAT	134000	
/notes="clone overlaps with GenBank Accession Number AC108896 clone RP42-89G3 (center project name axh)"		960	TATAGGGGACTGGAATGCAAAAGTAGGAAGCAAAACACCTGTGAGTAAACAGGCAAAAT	1019	
97016..154440		133999	TATAGGGGACTGGAATGCAAAAGTAGGAAGTCAAGAAACACCTGTGAGTAAACAGCAAAAT	133940	
/notes="assembly_fragment clone_end:SP6 vector_side:right"		1020	TGCGCTTGGAATACGGAATGAAGCAGGCAAGAGACTAATAGAGTTTTCGCCAAGAAAATGC	1079	
Query Match 52.9%; Score 1889.31; DB 1; Length 154440; Best Local Similarity 73.4%; Pred. No. 0; Matches 2526; Conservative 0; Mismatches 292; Indels 623; Gaps 15;		133939	TGCGCTTGGAAGTACAGATGAAGCAGGCAAGGCTAATAGAGTTTTCGCCAAGAGAACAC	133880	
1	GTCAGGAAGGGCGCAGTGAGGAGAGGTACCTACCTCGTCCAAAGTAGGAGCAGTAGC	60	1080	ACTGGTCTATAGCAAAACACCCCTCTTCCAAACACACAAAGAGAGACTCTACATGACATC	1139
134949	GTCAGTAGCAGCGCGCTGAGGAGA-----TACCCCATGTCCAAAGTAGGAGCAGCGGC	134895	133879	ACTGGTCTATAGCAAAACACCCCTCTTCCAAATAACACAAAGAGAGACTCTACATGACATC	133820
61	TGCGCTTTGCTGAGCAGCGCTTAAGAGATACCCAGCGCCCAAGGTAAAGAAACCCCAAG	120	1140	ACCAGATGGTCAACACCCGAAATCAGATTGATTATATTTCTTTGACGCCAAAAGTGGAGAG	1199
134894	TGTCCTTTGCTGAGCAGCGCTGAAGAGATACCCCAAGTCCAAAGTAAAGAAACCCCAAG	134835	133819	ACCAGATGGCCCAACACCAAAATCAG-----ATTCTTTGGAGCCAAAGATGGAGAGAG	133769
121	TAAGATGGTAGGTGTTGTGAGAGGGCATCAGAGGGCAGACATCTGAAACCATACACGCA	180	1200	CTCTATACAGTACAGCAAAAAACAAGACCCAGTGTGTGGCTCAGATCATGAACTCCT	1259
134834	TAAGATGGTAGGTGCTGAGAGAGGGGATCAGCGGCGACAGCGACTGAAAACCAATACAC	134775	133768	CTCTATATAGTACAGCAAAAAACAAGACCCGTTGCTGACTATGGCTCAGATCATGAACTCCT	133709
181	GAAACTAGTCAATCTAATCTACACTAGGACACAGCGCTTGTCTAACTCAATGAAACTAAG	240	1260	TATTCGCCAAATTCAGACTTTAAATTTGAAGAAAGTAGGGAAAACCACTAGATCACTCAGGTA	1319
134774	GAAACTATCCCAATCTGATCACAC--AGACCACAGCGCTTGTCTAACTTAATGAAACTAG	134717	133708	TATTCGCCAAATTCAGACTTTAAA--TGAAGAAAGTAGGGAAAACCACTAGACCAATTCAGGTA	133650
241	CCATGCGCTTGGGCAACCAAGATGGCGAGGTGATGTTGGGAGAGATCTGACAGAAATGTG	300	1320	AGACCTTAATCAATCCCTTATGATTATACAGTGAAGTGAAGTGAAGTATGATTTAAGGGCCT	1379
134716	CCATGCTGTGGGCCATCCAGACGGCGGCTCATGTTGGAG--GGTCTGACAGAAATGTG	134658	133649	TGACCTTAATCAATCCCTTACGATTATACAGTGAAGTGAAGATATATTTTAAGGGACT	133590
301	GTCACCTGGAAGAGGGAAT--GCAACCACTTCAGTATTTCTTGCTTGAGAACCCCATGAA	359	1380	AGATCTGATAGACAGAGTACCTTAATGAACCTATGGAACAGAGGTTTCATGACATTTGTACAGGA	1439
134657	GTCACCTGGAAGAGGGAATGGCAACCACTTCAGTATTTCTTGCTTTAGAACCCCATGAA	134598	133589	AGATCTGATAGACAGAGGCGCTGATGAACCTATGGAACAGAGGTTTCATGCAATTTGTACAGGA	133530
360	CAGTATGAAAGGCAAAATGATAGTACTGAAAGAGGAACTCCCGAGGTCAGTAGGTGC	419	1440	GACAGGGATCGAGA--CCATCCCATGGAAGAAAGTGAAGAAAGCAAAATGGCTGTCTG	1498
134597	CAGTATGAAAGGTAAAGAGCAGACCGCTGAAAGATGAACCTCCCGAGGTCGGTAGGTGC	134538	133529	GACAGGGATCAAAACCCATCCCATAGAAAAAATGCAAAAAAGCAAAATGGCTGTCTG	133470
420	CCCATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGAAATGAAGAGATGGAGCC	479	1499	GGGAGGCGCTTACAAATAGCTGTGAAAAAGAGAGAGTGAAGAAAGCAAAAGAAAAAGGAAA	1558
134537	CCCATGCTACTGGAGATCAGTGGAGAAACAACTCCAGAAAG--ATGAAGGGATGAAGCT	134479	133469	AGGAGGCTTTTAAATGGCTGTGAAAGAGAGAGAGCGAAAGCAAGGAGAAAGGAAA	133410
480	AAAGCAAAAAGATACCCAGCTGTGGATGTGATGTTGATATTAAGCAAGTCCGATGCTG	539	1559	GATAAAGCATCTGAATGACAGAGTTCCTAAA-----	1588
134478	AAAGCAAAAACACACCCAGTTGCGGATGTGATCTGGTATGAAGCAAGTCCGATGCTA	134419	133409	GATATAAACATCTGAATGCAGAGTTTCCAAAGAACACAGCAAGAGAGATAAGAAAGCCTTCT	133350
540	TAAAGAGCAATATTGCAATGAGAACCTGGAATGTCAAGTCCATGAATCAAGGCAAAATGGA	599	1589	-----	1588
134418	TAAAGAGCAATATTGCAATGAGAACCTGGAATGTGATGTTTCCATTAATCAAGGAAATTTGGA	134359	133349	TCAGCGATCAGTGCACAAAGCATAGAGGAAAACACAGAAATGGGAAAGACTAGAGATCTTT	133290
600	AGTGGTCAAAACAGAGATGCGAAGAGTGAATGTCAACATTTCTAGGAATCAGCGAACTAAA	659	1589	-----	1588
134358	AGTGGTCAAAACAGAGATGCGAAGAGCGAAACGTCACATTTCTAGGAATCAGTGAATAAG	134299	133289	TCAAGAAAATCAGAGATACCAAGGGAACATTTTCATGCAAGAGATGGGCTCAATAAAGGACA	133230
660	ATGGACTGGAATGGGTGAATTTAACTCAGATGACCAATTAATCTACTCTGCGGCGAGGA	719	1589	-----	1588
134298	ATGGACTGGAATGGGTGAATTTAACTCAGATGACCAATTAATCTACTCTACTCTGCGGCGAGGA	134239	133229	GAAATGGTATGGACCCCAACAGACAGAGAGATATTTAAGAGAGGTGGCAAGATACACAG	133170
720	ATCCCTCAGAGAAATGGAGTAGCCATCATGGTCAACAAAGAGTCCGAAATCAGTACT	779	1589	-----	1588
134238	ATCCCTTAGAAGAAATGGAGTAGCCATCATAGTCAACAAAGAGTCCCAAAATGCGTACT	134179	133169	GAGAACTGTACAAAAAAGATCTTTCATGCCAGAGATAATCAGATGGTGTGTGATCACTACC	133110
780	TGGATGCACTCTCAAAAACGACAGATGATCTCTGTTGTTTCCAGGCAAAACCATTCAA	839	1589	-----	1588
134178	TGGATGCAATCTCAAAAACGACAGAAATGATCTCTGTTGTTTCCAGGCAAAACCATTCAA	134119	133109	TAGAGCCAGACATCCTGGAATGTGAAGTCAAGTGGGCGCTTAGGAAGTAAACAACGAACA	133050
			1589	-----	1588

Db	133049	AAGCTAGTGGAGGTGATGGAATTCAGTTCAGCTATTTTCAAAATCCTGAAAGATGATGCTG	132990	Db	131973	TCAGCACAAAATAATGATGCTTTTGAACCTGTGTGTTGGAGAAAGACTCTTGAGAGTCTC	131914
Qy	1589	-----	1588	Qy	2463	TTGGAAGTCAAGGAGATCCAAACCAAGTCCATTTCTGAAGGAGATCAAGCCCTGGGATTTCTTT	2522
Db	132989	TGAAATATTTGCACTCAATATGCCAGCAAAATTTGGAAACTCAGCAGTGGCCACAGGACT	132930	Db	131913	ATGGACTACAGGAGATCCAAACCAAGTCCATTTCTGAAGGAGATCAGTCTGGGTCTCTTT	131854
Qy	1589	-----	1588	Qy	2523	GGAAGGAATGATGCTAAAGCTGAAACTCCAGTACTCTTGGCCACCTGATCAGAGAGCTGA	2582
Db	132929	GGAAAAGGTCACTTTTGGTCCAAATCCCAAGAAAGGCAATGTCAAGAATGCTCAAACT	132870	Db	131853	GGAAGGAATGATGCTAAAGCTGAAACTCCAAATCTTTTGGCCACCTCATGCAAGAGTTGA	131794
Qy	1589	-----	1588	Qy	2583	CTCACTGAAAGAACCTCTGATGCTGGAGGAGTATGGGGGACAGGAGGAGGACGACA	2642
Db	132869	ACTGCACAATTCACCTCATCTCACACGCTAGTAAAGTAAGCCTCAAAATTTCTCCAAGCCA	132810	Db	131793	CTCATGTGAAAGATCTCTGATTTCTGGGAAAGATTTGAGTAGGAGAGTAGGGATGACA	131734
Qy	1589	-----	1625	Qy	2643	GAGGATGAGTGGCTGGATGGCATCACTGATCTGATGGACCTGAGTCTGGGTGAACCTCT	2702
Db	132809	GGCTTCAGCAATACGTGAACCTTCAAGATGTTCAAGCTGGTTTGTAGAAAAGGCA	132750	Db	131733	GAGGATGAGTGGCTGGATGGCATCACTCACTCAATGAACATGATTCGAGTACACCTCG	131674
Qy	1626	GAGGAACAGAGACCAAAATTCACCAATCTCTGTATCATGGAAGAAAGCAAGAGATTCC	1685	Qy	2703	GGAGTTGCTGATGACAGGAGGCTGCTGCTGGCGGATTCATGGGTGCACAAAGAGTTG	2762
Db	132749	GAGGAACAAGAGATTTAAATTTGCCAACATCCGCTGGATCATGGAAGAAAGCAAGAGATTCC	132690	Db	131673	GGAGTTGCTGAT-GACAGGAGGCTGGCATGCTCAGTCCATGGATCACAAGAGTCCG	131615
Qy	1686	AGAAAACATCTATTTCTGCTTTATTTGACTATGCAAGGCTTTGACTGTGGGGTCA	1745	Qy	2763	GACAGACTGAGCAACTGAACCTGAACCTGACTGACTGAAACCTTAGTATTTATATTAC	2822
Db	132689	AGAAAACATGATTTCTGCTTTATTGACTATGCAAGGCTTTGACTGTGGATCAGA	132630	Db	131614	GACATGACTGAGCAACTGAACCTGAACCTGAACCTGAACCTGAACCTGAACCTGAACCT	131555
Qy	1746	ATAAACTGTGAAAATTCGAAAGGATGGGAATACCAGACCACTGACCTGCTCTTGA	1805	Qy	2823	TCAGAAAATAGTATTTTCATA	2843
Db	132629	ATAAATGTGGAATTTCTCAAGAGTTGGGAATACCATTCCATCTGACCTGCTCTTGA	132570	Db	131554	TCAGGAAATTAATGACAGA	131534
Qy	1806	AAAATTTGATGAGGTGAGGAACCAACAGTAGTAACCTGACATGCAACAGACTGGT	1865	RESULT 39			
Db	132569	TGAATCTGTATGCAAGTCAAGAACCAACAGTAGTAACCTGACATGGAACCAAGAGCTGGC	132510	AC137647/c			
Qy	1866	TCCAAGTAGGAAAAGGAGTATGTCAGGCTGTATATTGTCAACCGGCTGTTTAACTTCT	1925	AC137647	165992 bp	DNA	linear
Db	132509	TCTAAATTTGGAAGGAGTACGTCAAGCTGTATATTGTCAACCTGCTTATTAACTTAT	132450	LOCUS	Bos taurus clone RP42-127F18,	WORKING DRAFT SEQUENCE,	6 ordered
Qy	1926	ATGCAGAG-ACATCATGAGAAACCTCGGCTGGAGGAAGCAACAGCTGGAAATCAAGATTG	1984	DEFINITION	pieces.		
Db	132449	CTGCAGAGCACATCATGAGAAA--CTGGGCTGGAGGAAGCAACAGCTGGAAATCAAGACTG	132392	ACCESSION	AC137647.2	GI:28201542	
Qy	1985	CCGGGAGGAATGAGCAATACCTCAGATATGAGATATACACCTTATGCGAGAAAGTG	2044	VERSION	HTG; HTGS_PHASE2; HTGS_DRAFT.		
Db	132391	CCTAGAGAAATATCAGTAACCTCAGATACGAGATGACACACCTTATTGCGAGAGTTA	132332	KEYWORDS	Bos taurus	(cow)	
Qy	2045	AAGAGGAACATAAAGCCCTCTTGATGAAGGTGAAAGGAGAGTGAAAAAGTTGGCTTAA	2104	SOURCE	Bos taurus		
Db	132331	AAGAAGAACTAAAGAGCCCTCTTGATGAAGGGAAGGAAGGAAGTAAAGAAAGTTGGCTGAA	132272	ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
Qy	2105	AGCTCAACATTCAGAAACGAGATCATGCTGCTGCCATCCTCATCTGAGGAATAG	2164	REFERENCE	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;		
Db	132271	AGCTCAATATTCAAGAAATGAGATCATGCGATCTGCTCCATCACTTCATGGCAAAATAG	132212	AUTHORS	Bovidae; Bovinae; Bos.		
Qy	2165	ATGGGAAACAGTGAACAGCTGTCAGACTTTATTTTGGGGGCTCCAAAATCACTGC	2224		1 (bases 1 to 165992)		
Db	132211	ATGGGAAATAGTCGAACAGTGGCTGACTTTATTTCT--TGGGCTCCAAAATCACTGC	132154	JOURNAL	Akhter N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,		
Qy	2225	AGATGGTGAAGTCAAGCAATGAAATTAAGACACTTACTCTCTGGAGAAAGTTA--AC	2282	REFERENCE	Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,		
Db	132153	AGATGGTGAAGTCAAGCAATGAAATTAAGACACTTACTCTCTTGAAGGAAGTTATGAT	132094	AUTHORS	Carliaga, K., Coleman, S., Engle, J., Granite, S., Guan, X., Gupta, J.,		
Qy	2283	CAACTAGATAGCATTTGAAAGCAGACATTAACCTTGCCACAAAGCCCACTCTAGT	2342	JOURNAL	Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E.,		
Db	132093	CAACTAGATAGCATTTGAAAGCAGACATTAACCTTGCCACAAAGTCCGCTCTAGT	132034	REFERENCE	Laric, P., Lee-Lin, S.-Q., Legaspi, R., Maduro, Q.B., Maduro, V.B.,		
Qy	2343	CAAGCTATGTTTTTCCAGTGGTATGATGATGATGAGATTGGACTGTGAAGAAAGC	2402	TITLE	Marquies, E.H., Masello, C., Maskeri, B., McDowell, J.,		
Db	132033	CAAGCTATGTTTTTCCAGTGGTATGATGATGATGATGATGATGATGATGATGATGATGAT	131974	JOURNAL	Paguirigan, C., Pearson, R., Portnoy, M.E., Prasad, A.,		
Qy	2403	TGAGCACTGAAGATTGATGCTTTTGAACCTGTTGTTGAGAGAGACTTTGAGAGTCCC	2462	AUTHORS	Reddix-Dugue, N., Schandler, K., Schueler, M.G., Sisor, C.,		
					Stantripo, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L.,		
					Wetherby, K.D., Wiggins, J., Young, A. and Green, E.D.		
					NISC Comparative Sequencing Initiative		
					Unpublished		
					2 (bases 1 to 165992)		
					Green, E.D.		
					Direct Submission		
					Submitted (27-NOV-2002) NIH Intramural Sequencing Center, 8717		
					Grovermont Circle, Gaithersburg, MD 20877, USA		
					3 (bases 1 to 165992)		
					Green, E.D.		
					Direct Submission		
					Submitted (04-FEB-2003) NIH Intramural Sequencing Center, 8717		
					Grovermont Circle, Gaithersburg, MD 20877, USA		
					On Feb 4, 2003 this sequence version replaced gi:25699231.		

					Genome Center		
					Center: NIH Intramural Sequencing Center		
					Center code: NISC		
					Web site: http://www.nisc.nih.gov		
					Contact: nisc_zoo@nhgri.nih.gov		

----- Project Information
Center project name: crr
Center clone name: 127F18

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig, has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 165185 bases at least Q40
Consensus quality: 165337 bases at least Q30
Consensus quality: 165454 bases at least Q20
Insert size: 155000; agarose-fp
Insert size: 165492; sum-of-contigs
Quality coverage: 11.30x in Q20 bases; agarose-fp
Quality coverage: 10.58x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 6 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 58794: contig of 58794 bp in length
58795 58894: gap of unknown length
58895 84062: contig of 25168 bp in length
84063 84163: gap of unknown length
84163 100779: contig of 16617 bp in length
100780 100879: gap of unknown length
100880 133011: contig of 32132 bp in length
133012 133112: gap of unknown length
133112 144769: contig of 11658 bp in length
144770 144869: gap of unknown length
144870 165992: contig of 21123 bp in length.

FEATURES
source

1..165992
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="RP42-127F18"
/clone_lib="RP42"
1..58794
/note="assembly_fragment
clone_end:17
vector_side:left"
58895..84062
/note="assembly_fragment"
84163..100779
/note="assembly_fragment"
100880..133011
/note="assembly_fragment"
114737..165992
/note="clone overlaps with GenBank Accession Number AC137906 clone RP42-73K10 (center project name crr)"
133112..144769
/note="assembly_fragment"
144870..165992
/note="assembly_fragment
clone_end:SP6

vector_side:right

Query Match	52.8%;	Score 1887;	DB 1;	Length 165992;
Best Local Similarity	75.8%;	Pred. No. 0;		
Matches 2335;	Conservative	0;	Mismatches 145;	Indels 602; Gaps 6;
QY	315	GGAAATGCAAAACACATTCAGTATTTCTTGCTTGAGAACCCCATGACAGTATGAAAGGCA	374	
Db	69368	GCAATTAGAAATACATTCAGTACTCTTGCTTGAGAACCCCATGAATAATATGAAAGGCA	69309	
QY	375	AAATGATAGGATACATGAAAGAGGAATCTCCAGGTGAGTGTGCCCCATATGCTACTGG	434	
Db	69308	AAATGATAGGATACATGAAAGAGGAATCTCCAGGTGAGTGTGCCCCATATGCTACTGG	69249	
QY	435	AGATCAGTGGAGAAATAACTCCAGAAAGAAATGAAGAGATGAGCCAAAGCAAAAAGAAATA	494	
Db	69248	AGATCAGTGGAGAAATAACTCCAGAAAGAAATGAAGAGATGAGCCAAAGCAAAAAGAAATA	69189	
QY	495	CCAGCTGTGGATGTGACTGTGATATAGCAAGGTCCGATGCTGTAAAGAGCAATATTG	554	
Db	69188	CCAGATGTAGATGTGACTGTGATAGAAAGAGGTCCGATGCTGTAAAGAGCAATATTG	69129	
QY	555	CATAGAACCTTGGATGTGAGTCCATGAATCAAGGCAAAATGGAAGTGTCAACAAGA	614	
Db	69128	CATAGAACCTTGGATGTGAGTCCATGAATCAAGGCAAAATGGAAGTGTGTCAACAAGA	69069	
QY	615	GATGCAAGAGTGAATGTCAACATTTCTAGGAATCAGCGAATCTAAATGAGCTGGATGG	674	
Db	69068	GATGCAAGAGTGAATGTGACATTTCTAGGAATCAGTGAATGAGCTGGATGG	69009	
QY	675	TGAATTTACTCAGATGACCAATATATCTACTGCGGAGGAGGATCCCTCAGAGAA	734	
Db	69008	TGAATTTACTCAGATGACCAATATATCTACTGCGGAGGAGGATCCCTCAGAGAA	68949	
QY	735	TGGAGTAGCATCATGCTCAACAAAGAGTCCGAATGCGAGTCTGGATGAGTCTCAA	794	
Db	68948	TGGGTGCGCATCATGCTCAACAAAGAGTCCGAATGCGAGTCTGGATGAGTCTCAA	68889	
QY	795	AAACGACAGATGATCTCTGTTTGTTCAGAGGCAAAACCAATCAATATCACAGTAAATCA	854	
Db	68888	AAACGACAGATGATCTCTGTTTGTTCAGAGGCAAAACCAATCAATATCATGTAATCCA	68829	
QY	855	ACTCTATGCCCCCAACCAAGTAAATGCTGAAGAGCTGAAGTGAACGCTCTTATGAAGACT	914	
Db	68828	AGTCTATGCCCCCAACCAAGTAAATGCTGAAGAGCTGAAGTGAAGTCTTATGAAGACT	88770	
QY	915	ACAAGACCTTTTAGAATCAACCCCAAAAGAGTCTCTCTCATTTATAGGGGACTGGAA	974	
Db	68769	ACAAGACCTTTTAGAATCAACCCCAAAAGAGTCTCTCTCATTTATAGGGGACTGGAA	68710	
QY	975	TGCAAAAGTAGGAAGCAAAAGAACCTTGGAGTACAGGCAAAATTTGGCTTGGATACG	1034	
Db	68709	TGCAAAAGTAGGAAGTAAAGAAACACCTTGGAGTAAAGCAAAATTTGGCTTGGATACG	68650	
QY	1035	GAATGAAGCAGGCGCAAAAGACTAATAGATTTTTCGCAAGAAATATGCACTGGTCTATAGCAAA	1094	
Db	68649	GAATGAAGCAGGCGCAAAAGACTAATAGATTTTTCGCAAGAAATATGCACTGGTCTATAGCAAA	68590	
QY	1095	CACCTCTTCCAAACACACAGAGAGACTCTTACATGAGCATCACAGATGGTCAACA	1154	
Db	68589	CACCTCTTCCAAACACACAGAGAGACTCTTATACATGAGCATCACAGATGGTCAACA	68530	
QY	1155	CCGAATCAGATTTGATTTATTTCTTTTGGAGCCAAAGATGGAGAGCTCTATACAGTACG	1214	
Db	68529	CTGAAATCAGATTTGATTTATTTCTTTTGGAGCCAAAGATGGAGAGCTCTATACAGTACG	68470	
QY	1215	AAAAACAAGCAGGAGCTTACTGTGGTTCAGATCATGAATCTCTTATTCGCAAAATTCAG	1274	
Db	68469	AAAAACAAGCAGGAGCTTACTGTGGTTCAGATCATGAATCTCTTATTCGCAAAATTCAG	68410	
QY	1275	ACTTAATTTGAAGAGTAGGGAACCACTAGATCACTCAGGTAGAGCTAAATCCCAT	1334	
Db	68409	ACTTAATTTGAAGAGTAGGGAACCACTAGATCACTCAGGTAGAGCTAAATCCCAT	68350	

1335 CCCTTATGATTATACAGTGGAAAGTGAAGATAGATTTAAGGSCCTAGATCTGTATAGACAG 1394
Db 68349 CCCTTATGATTATACAGTGGAAAGTGAAGATAGATTTAAGGSCCTAGATCTGTATAGATAG 68290
1395 AGTACCTAATGAATATGAGACAGAGGTTTCAATGACATTTGACAGAGACAGAGGATCGAGAC 1454
Db 68289 AGTGCAGATGAATATGGAATGAGGTTTGGTGAATTTGACATTTGACAGAGACAGAGGATCAAGAC 68230
1455 CATCCCATGAAAGAAATGCAAAAGCAAAATGAGTGTCTGCTGGGAGGCTTACAAAT 1514
Db 68229 CATCCCATGAAAGAAATGCAAAAGCAAAATGAGTGTCTGCTGGGAGGCTTACAAAT 68170
1515 AGCTGTGAAAAGAGAGAAAGTGAAGCAAAAGGAAAAAGAAAGATATAAAGCATCTGAA 1574
Db 68169 AGCTGTGAAAAGAGAGAAAGTGAAGCAAAAGGAAAAAGAAAGATATAAAGCATCTGAA 68110
1575 TGCAGAGTTCCAAA----- 1588
Db 68109 TACAAAGTTCCAAAAGATAGCAAGAGAGATAGAAAAGCCTTTCTTACGCGATCAATGCAA 68050
1589 ----- 1588
Db 68049 AGAATAGAGAAAAACAACAGATGGGAAGACTAGGGATCTCTTCAAGAAATCAGAGA 67990
1589 ----- 1588
Db 67989 TACCAAGGAACATTTTCATGCAAGATGAGTCGATTAAGGACAGAAATGGTATGGACCT 67930
1589 ----- 1588
Db 67929 AACAGAGCAGAGATATTAAGAGAGGTGGCAAGAAATACACAGAGAAATTTGACAAAAA 67870
1589 ----- 1588
Db 67869 AGATCTTACAGACCCAGATATCATCATGTTGATCACTGACCTAGAGCCAGACATCCT 67810
1589 ----- 1588
Db 67809 GGAATGTGAAGTCAACTGGGCTTTAGAGGATCACTATGAAAGCTAGTGGAGGTGA 67750
1589 ----- 1588
67749 TGGAAATCCAGTTGAGCTATTCCAAATTCCTGAAAGATGATGCTGTGAAAGTCTTCACTC 67690
1589 ----- 1588
67689 AATATGCCAGCAAAATTTGGAAGTCTAGAGTGGCCACAGGACTGGAAAAAGGTCAAGTTT 67630
1589 ----- 1588
67629 CATTCCAATCCCAAGAAAGGCTATGCCAAGAAATGCTCAAACTACCGCAAAATGCACT 67570
1589 ----- 1588
67569 CATCTCAAGCTAGTAAAGTAATGCTCAAAATTCCTCAAGCCAGCCTTCAGCAATATGT 67510
1589 -----GAACTTCCAGTTGTTCAAGCTGTTTNGAAGAGTCAAGAGACCAAGACCA 1641
67509 GAACCGTGAACCTTCTGATGTTCAAGCTGTTTGTAGAAAAGGAGAGACCAAGACCA 67450
1642 AATTGCCAACATCTCTGTATCATGGAAGAAAGCAAGAGTTCAGAAAAAATCTATTT 1701
Db 67449 AATTGCCAACATCTCTGTGATCATGGAAGAAAGCAAGAGTTCAGAAAAAATCTATTT 67390
1702 CTGCTTTATTGACTATGCAAGGCTTTGACTGTGGGGTCAAAATTAAGCTGTGGAAAT 1761
Db 67389 CTGCTTTATTGACTATGCAAGGCTTTGACTGTGGGGTCAAAATTAAGCTGTGGAAAT 67330
1762 TCTGAAAGGATGGGAATACCAAGACCACTGACTGCTGCTTCAAAATTTGATGCAGG 1821
67329 TCTGAAAGAGATGGGAATACCAAGACCACTGACTGCTTCAAGAAATTTGATGCAGG 67270

1822 TCAGGAAGCAACAGTTAGAACTGGACATGGAAACAACAGACTGGTTCCAAAGTAGGAAAAAGG 1881
Db 67269 TCAGGAAGCAACAGTTAGAACTGGACATGGAAACAACAGACTGGTTCCAAATAGGAAAAAGG 67210
1882 AGTATGTCAAGCTGTATATTGTCAACCGGCTTTTAACTTCTATGAGAG-ACATCAT 1940
Db 67209 AGTTCGTCAAGCTGTATATTGTCAACCGCTTTTAACTTCTATGAGAGTATATCAT 67150
1941 GAGAAACGCTGGCTGGGAAGAGCAACAGCTGGAATCAAGATTGCGGAGAGAAATAGCAA 2000
Db 67149 GAGAAACACTGGAAGTGAAGCAACAACTGAATCAAGATTGCGGAGAGAAATATCAA 67090
2001 TAACTCTAGATATGAGATGATATACACCCCTTATGCGAGAAAGTGAAGAGGAACATAAAG 2060
Db 67089 TAACTCTAGATATGAGATGATATACACCCCTTATGCGAGAAAGTGAAGAGGAACATAAAG 67030
2061 CCTCTGATGAGTGAAGAGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAA 2120
Db 67029 CCTCTGATGAAAGTGAAGTGAAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAA 66970
2121 AACGAAGATCATGGCATCTGGTCCCATCACTTCATGCGAAATAGATGGGAAAACAGTGGGA 2180
Db 66969 AACGAAGATCATGGCATCTGGTCCCATCACTTCATGCGAAATAGATGGGAAAACAGTGGGA 66910
2181 AACAGTGCAGACTTTATTTTGGGGGCTCCAAATCACTGCAGATGGTGAAGTGCAGC 2240
Db 66909 AACAGTGCAGACTTTA--TTTGGGGGTTTCCAAATCACTGCAGATGGTGAAGTGCAGC 66852
2241 CATGAAATTTAAAGACACTTACTCTCTGGAAGAAAGTTA--ACCAACTAGATAGCATA 2298
Db 66851 CATGAAATTTAAAGACACTTACTCTCTGGAAGAAAGTTA--ACCAACTAGATAGCATA 66792
2299 TTGAAAGCAGAGACATTAACCTTCCAAACAAAGCCCATCTAGTCAAGGCTATGGTTTT 2358
Db 66791 TTCAAAGCAGAGACATTAACCTTCCAAACAAAGGTTCTGTAGTCAAGGCTATGGTTTT 66732
2359 CCAGTGGTCACTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2418
Db 66731 CCTGTGGTCACTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 66672
2419 GATGCTTTGAACTGTGGTGTGAGAGAGACTCTTTGAGAGTCCCTTGGACTGCAAGGAGA 2478
Db 66671 GATGCTTTGAACTGTGGTGTGAGAGACTCTTTGAGAGTCCCTTGGAGTCAAGGAGA 66612
2479 TCCAAACAGTCCATTCAGAGAGATCAGCCCTGGGATTTCTTTGGAAGGATGATGCTA 2538
Db 66611 TCCAAACAGTCCATTCAGAGAGATCAGCCCTGGGATTTCTTTGGAAGGATGATGCTA 66552
2539 AAGCTGAAACTCCAGTACTTTGGCCACCTGATCAGAGAGCTGACTCACTGGAAGAGACC 2598
Db 66551 AAGCTGAAACTCCAGTACTTTGGCCACCTGATCAGAGAGTGTGACTCATTTGGAAGAGACT 66492
2599 CTGATGCTGGAGAGGATTTGGGGCAGAGAGAGAGAGAGAGAGAGAGAGAGATGAGTGGCTG 2658
Db 66491 CTGATGCTGGAGAGGATTTGGGGCAGAGAGAGAGAGAGAGAGAGAGAGATGAGTGGCTG 66432
2659 GATGCTCACTGACTCGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2718
Db 66431 GATGCTCACTGACTCGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 66372
2719 AGGAGGCTGTCTGCGGAGATTCATGGGCTCAAAAGAGTGTGAGCAGCAGTCAAGCAAC 2778
Db 66371 AGGAGGCTGTCTGCGGAGATTCATGGGCTCAAAAGAGTGTGAGCAGCAGTCAAGCAAC 66315
2779 TGAACTGAACTGAACTGACTG 2800
Db 66314 TGAACTGAACTGAACTGACTG 66293

RESULT 40

AC129960/c

LOCUS

DEFINITION

AC129960 190441 bp DNA linear HTG 25-JAN-2003
Bos taurus clone RP42-362K14, WORKING DRAFT SEQUENCE.

ACCESSION

AC129960.2 GI:27901861
 HTG; HTGS PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
 SOURCE
 ORGANISM

Bos taurus
 Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 190441)
 Bos taurus, clone RP42-362K14

REFERENCE

1 (bases 1 to 190441)

AUTHORS

Unpublished

JOURNAL

REFERENCE

AUTHORS

Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
 Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B.,
 Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
 Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
 Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,
 Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,
 Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
 Karatas, A., Kellis, C., Landers, T., Levine, R., Lindblad-Toh, K.,
 Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,
 McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mienga, V.,
 Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H.,
 O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
 Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P.,
 Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
 Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
 Zembek, L., Zimmer, A. and Zody, M.

TITLE

JOURNAL

Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

AUTHORS

3 (bases 1 to 190441)
 Birren, B., Nusbaum, C., Lander, E., Abouelheil, A., Allen, N.,
 Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T.,
 Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
 Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K.,
 Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
 Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,
 Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
 Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kellis, C., Landers, T., Levine, R.,
 Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C.,
 Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
 Meldrum, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J.,
 Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
 Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C.,
 Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
 Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,
 Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

JOURNAL

Submitted (25-JAN-2003) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jan 25, 2003 this sequence version replaced gi:22123282.

All repeats were identified using RepeatMasker:

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

COMMENT

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L24461

Center clone name: 362_K.14

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 190436 bases at least Q40

Consensus quality: 190441 bases at least Q30

Consensus quality: 190441 bases at least Q20

Insert size: 198000; agarose-fp

Insert size: 190441; sum-of-contigs

Quality coverage: 16.1 in Q20 bases; agarose-fp

Quality coverage: 16.8 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently

* consists of 1 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have

* provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 190441: contig of 190441 bp in length.

assembly_fragment.

location/Qualifiers

source

1. 190441

/organism="Bos taurus"

/mol_type="genomic DNA"

/db_xref="taxon:9913"

/clone="RP42-362K14"

/clone_lib="RPC1-42 Bovine BAC"

Query Match

Best Local Similarity 52.8%; Score 1885.1; DB 1; Length 190441;

Matches 2467; Conservative 0; Mismatches 254; Indels 619; Gaps 12;

QY 62 GGCCTTCTGAGCAGCGGTAAAGAGATACCCACGCCCAAGTAAGAGAACCAAGT 121

Db 137961 GGCCTTCTGAGTAGCGGTGAAGAGATACCCACGCCCAAGTAAGAGAACCAAGT 137902

QY 122 AAGATGTTAGTGTGTGTGAGAGGGATCAGAGGGGAGAGATCTGAAACCAATACAGCAG 181

Db 137901 AAGACAGTAGGCACTGAGAGAGGGATCAGAGGGGAGAGATCTGAAACCAATACAGCAG 137842

QY 182 AAACTAGTCAATCTAATCACTAGGACCAAGCCTTGTCTAATCTAATGAACTAAGC 241

Db 137841 ACAACAATCTGAT-----CACATGGACCAAGCTGTCTAATGAACTAAGCCTTGTGT 137789

QY 242 CATGCCGTGGGGCAACCAAGATGGCAGGTCATGTTGGAGAGATCTGACAGAAATGTTGG 301

Db 137788 GGGGGCAACCCAG--ACCCACATGTTGGGTCTGTTGGAGAGTCTGACAGAAATGTTGG 137731

QY 302 TCCACTGGAGAGGGAAT-GCAAAACCACTTCAGTATTTTGCCTTGAGAACCCCAATGAAC 360

Db 137730 TCCACTGAAGAGGGAATGGCGAACCACTTCAGTATTTTGCCTTGAGAACCCCAATGAAC 137671

QY 361 AGTATGAAAGGCAAAATGATAGTACTGAAAGGGAACCTCCCAAGGTCTAGTAGTGCC 420

Db 137670 AGTATGAAAGGCAAAATGATAGTACTGAAAGGGAACCTCCCAAGGTCTAGTAGTGCC 137611

QY 421 CCATATGTTACTGGAGATCAGTGGAGAAATACTCCAGAAAGATGAAGAGATGAGGCCA 480

Db 137610 CAATATGTTACTGGAGATCAGTGGAGAAATACTCCAGAAAGATGAAGAGATGAGGCCA 137551

QY 481 AAGCAAAAAGATACCCAGCTGTGGATGTGACTGGTGTATTAAGCAAGGTCCGATGCTGT 540

Db 137550 AAGCAAAAAGATACCCAGCTGTGGATGTGACTGGTGTATTAAGCAAGGTCCGATGCTGT 137491

QY 541 AAAGAGCAATATTGCAATAGGAACCTGGAATGTCCAGTCCATGAATCAAGGCAAAATTGAA 600

Db 137490 AAAGAGCAATATTGCAATAGGAACCTGGAATGTCCAGTCCATGAATCAAGGCAAAATTGAA 137431

QY 601 GTGGTCAAAACAGAGATGGCAAGAGTGAATGTCAACATTTCTAGGAATCAGGGAATCAAAA 660

Db 137430 GTGGTCAAAACAGAGATGGCAAGAGTGAATGTCAACATTTCTAGGAATCAGTGAATCAAAA 137371

QY 661 TGGACTGGNATGGGTGAATTTAACTCAGATGACCATTATATCTACTTCTGCGGGCAGAA 720

Grovemont Circle, Gaithersburg, MD 20877, USA
 On Dec 5, 2002 this sequence version replaced gi:18581925.
 ----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: <http://www.nisc.nih.gov>
 Contact: nisczoo@nhgri.nih.gov
 ----- Project Information
 Center project name: cjr
 Center clone name: 136015

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

```

----- Summary statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 154307 bases at least Q40
Consensus quality: 154347 bases at least Q30
Consensus quality: 154358 bases at least Q20
Insert size: 121000; agarose-1p
Insert size: 154377; sum-of-contrigs
Quality coverage: 12.74x in Q20 bases; agarose-1p
Quality coverage: 9.99x in Q20 bases; sum-of-contrigs

```

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * been provided by the submitter.

```

* This sequence will be replaced
* By the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 88559: contig of 88559 bp in length
* 88560: gap of unknown length
* 88561: contig of 88561 bp in length
* 126656: contig of 126656 bp in length
* 126657: gap of unknown length
* 126658: contig of 126658 bp in length.

```

FEATURES source

```
1. 154577
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="RP42-13c015"
/clone_lib="RP42"
1. 88559
/clone="assembly_fragment"
/clone_end:SP6
vector_side:left"
1. 14623
/clone="clone overlaps with
AC109792 clone RP42-475P
5711". 154577
/clone="clone overlaps with
AC109798 clone RP42-71E1
88660". 126656
/clone="assembly_fragment"
126757. 154577
/clone="assembly_fragment"
clone_end:T7
vector_side:right"
```

misc feature

misc feature

misc feature

misc feature

misc feature

Qy	2284	AACTAGATAGCATATTTGAAAGAGAGACATTTACCTTGCACAAAGCCCACTTAGTC	2343
Db	135159	AACCTAGACGATATTTAAAAAGCAGACATTTACTTTGTCAAAAAGATCCCGCTAGTT	135100
Qy	2344	AAGGCTATGGTTTTTCCAGTGGTCAATGATCGGATGTGAGAGTTGGACTGTGAAGAAAGCT	2403
Db	135099	AAGGCTATGGTTTTTCCAGTAGTCATGATGAGATGTGAGAGTTGGACTATAAAGAAAGCT	135040
Qy	2404	GAGCATGTAAGAATTTGATGCTTTTGAACATGTGGTGTGAGAGAGACTTTTGAGAGTCCCT	2463
Db	135039	GAGCGCT--GAATTTGATGTTTTTGAATTTGGTATTGGAAGACACTCTTTGAGAGTCACT	134983
Qy	2464	TGAGCTGCAAGGAGATCCAAACAGTCCATCTCTGAAGGAGATCAGCCCTGGGATTTCTTG	2523
Db	134982	TGAGCTGCAAGGAGATCCAAACAGTCCATCTTAAGAGGATCAGTCTCTGGTGTTCATGG	134923
Qy	2524	GAAGGAATGATGCTAAAGCTGAAACTCCAGTACTTTTGGCCACTGTATCAGAAGAGCTGAC	2583
Db	134922	GAAGGACTGATGTTGAAGCTGAAACTCCAATACTTTGATCACTTCAGGCGAAGAGTTGAC	134863
Qy	2584	TCAGCTGAAAAGACCTCGATGCTGGAGGGATTTGGGGCAGGAGGAGAAGGGGACACAG	2643
Db	134862	TCATTTGAAAAGACTCTGATGCTGGAGGGATTTGGGGCAGGAGGAGAAGGGGACGACAG	134803
Qy	2644	AGGATGAGATGGCTGGATGGCATCACTGACTCGATGGAAGTGAATCTTGGGTGAACTTCCTG	2703
Db	134802	AGGATGAGATGGCTGGATGGCATCACTGACTCAATGGAAGTGAATCTCAGTGAACCTCCGG	134743
Qy	2704	GAGTTGGTGAATGGACAGGGAGGCGCTGTCCTCGGGCGATTTCATGGGGTGCACAAAGAGTTGG	2763
Db	134742	GAGTTGAATGATGACAGGGAGGCGCTGGCGTCTGGGATTCATGGGGTGCACAAAGAGTTGG	134683
Qy	2764	ACACGACTGAGCAACTGAACCTGAACTGAACCTGTAAGAA	2803
Db	134682	ACACGACTGAGCAACTGAACCTGAACTGAGCGCCCTGGAAGAA	134643

RESULT 41	AC109787	154577 bp	DNA	linear	HTG 05-DEC-2002
AC109787	Bos taurus clone RP42-136015,	WORKING DRAFT SEQUENCE, 3 ordered			
LOCUS					
DEFINITION					
	pieces.				
ACCESSION	AC109787				
VERSION	AC109787.2	GI:36050954			
KEYWORDS	HTG; HTGS_PHASE2; HTGS_DRAFT.				
SOURCE	Bos taurus (cow)				
ORGANISM	Bos taurus				
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;				
	Bovidae; Bovinae; Bos.				
PREFERENCE	1 (bases 1 to 154577)				

REFERENCE	AUTHORS	TITLE
1 (Cases 1 to 154577)	Doenges, D. H.	NSRC Comparative Sequencing Initiative
	Antonicelli, A., Ayele, K., Beckstrom-Sternberg, S. M., Akhter, N., Blakesley, R. W., Bouffard, G. G., Brinkley, C., Brooks, S. Benjamin, B., Carriaga, C., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J. R., Karlins, E., Lalic, P., Lee-Lin, S.-Q., Legaspi, R., Maduro, Q. L., Maduro, V. B., Margulies, E. H., Masiello, C., Maskeri, B., McDowell, J., Peguigian, C., Pearson, R., Portuncu, M. E., Prasad, A., C., Reddy-Dugue, N., Schandler, K., Schueler, M. G., Sison, C., Stanirrop, S., Thomas, J. W., Thomas, P. J., Touchman, J. W., Vogt, J. L., Wetherby, K. D., Wiggins, L., Young, A. and Green, E. D.	

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE

ADAMSON
TITLE
Direct Submission
Submitted (05-DEC-2002) NIH Intramural Sequencing Center, 8717
Green, E.D.

Query Match 52.3%; Score 1868.9; DB 1; Length 154577;
Best Local Similarity 75.7%; Pred. No. 0;
Matches 2363; Conservative 0; Mismatches 141; Indels 619; Gaps 10;

QY 266 GGGCAGGTCATGGTGGAGATCTGCACAGATATGGTCCACTGGAGAGGGAAT-GCAAA 324
Db GGGCGGGTCATGGTGGAGAGCTCTGACAGATATGGTCCACTGGAGAGGGAATGGCAAA 59351

QY 325 CCACCTCAGTATTTCTTCCCTTGAGAACCCCATGACACAGATGATAAAGCGCAAAATGATAGG 384
Db CCACCTCAGTATTTCTTACCTTTGAGAACCCCATGACACAGATGATAAAGCGCAAAATGATAGG 59411

QY 385 ATACTGAAAGAGGAACCTCCACAGTCTCAGTAGTCCCATATGCTACTGAGAGATCAGTGG 444
Db ATACTGAAAGAGGAACCTCCACAGTCTCAGTCTGAGTCCCATATGCTACTGAGATCAGTGG 59471

QY 445 AGAAATTAATCCAGAAAGATGAAGAGATGGAGCCAAAGCAAAAGAAATACCCAGCTGTG 504
Db AGAAATTAATCCCAAAAGATGAAGGATGAAGCCAAAGCAAAATCAATCTCAGCTGTG 59531

QY 505 GATGTGACTGTGTATATAAGCAAGTCCGATGCTGTAAAGAGCAATATTCATAGGAACC 564
Db GATGTGACTGTGTATAGNAGCAAGTCCGATGCTGTAAAGAGCAATATTCATAGGAAC 59591

QY 565 TGGAAATGTCAAGTCCATGAATCAAGGCAAAATTTGGAAGTGTCAAAAGAGATGGCAAGA 624
Db TGGAAATGTGAGTCCATGAATCAAGGCAAAATTTGAGTGTCAAAAGAGATGGCAAGA 59651

QY 625 GTGAATGTCAACATCTAGGAATCAGCGAATCAAAATGGAATGGATGGGCTGAATTTAAC 684
Db GTCAACATCGACATCTAGGAATCAGCGAATCAAAATGGAATGGGCTGAATTTAAC 59711

QY 685 TCAGATGACCAATATATCTACTGCGGCGAGGAATCCCTCAGAGAAATGGATGAGTAGCC 744
Db TCAGATGACCAATATATCTACTGCTGGCGAGGAACCCCTTAGAGAAATGGATGAGTAGCC 59771

QY 745 ATCATGTCAACAAAGAGTCGGAATTCAGTACTTGGATGCAAGTCTCAAAAACGACAGA 804
Db ATCATGTCAACAAAGAGTCGGAATTCAGTACTTGGATGCAAGTCTCAAAAACGACAGA 59831

QY 805 ATGATCTCTGTTGTTTCCAAAGGCAAAACCATTAATCAATCAAGTATCCAAAGTCTATGCC 864
Db ATGATCTCTGTTGTTTCCAAAGGCAAAACCATTAATCAATCAAGTATCCAAAGTCTATGCC 59891

QY 865 CCAACCAAGTATGCTGAAGAGCTGAAGTTGAACGGTCTTATGAAGACCTTACAAAGACCTT 924
Db CCAACCAATATGCTGAAGAGCTGAAGTTGAACGGTCTTATGAAGACCTTACAAAGACCTT 59951

QY 925 TTGAACCTTAACACCCCAAAAGAGTCTCTTCTCATTTATGGGAGCTGGATGCAAAAGTA 984
Db TTGAACCTTAACACCCCAAAAGAGTCTCTTCTCATTTATGGGAGTGGATGCAAAAGTA 60011

QY 985 GGAAGCAAGAAACACCTGGAGTAACAGGCAAAATTTGGCCCTTGAATACGGAATGAAGCA 1044
Db GGACCTCAAGAAACACATGAGTAACAGGCAAAATTTGGCCCTTGAATACGGAATGAAGCA 60071

QY 1045 GGGCAAGACTATATAGATTTTGGCAAGAAATGCACTGTCTATAGCAACACCTCTTTC 1104
Db AGGCAAGAGCTATATAGATTTTGGCAAGAAATGCACTGTCTATAGCAACACCTCTTTC 60131

QY 1105 CAACAAACAAAGAGAACTCTACATGGACATCACAGATGGTCAACACCCGAAATCAG 1164
Db CACCAACAAAGAGAACTCTACATGGACATCACAGATGGTCAACACCCGAAATCAG 60189

QY 1165 ATTGATTTATTTCTTTCAGGCAAGATGAGAGCTCTATAGCTCAGCAAAACAAAGA 1224
Db ATTGATTTATTTCTTTCAGGCAAGATGAGAGCTCTATAGCTCAGCAAAACAAAGA 60247

QY 1225 CCAGGAGCTTACTGTGGCTCAGATCATGAATCTCTTATTCGCAAAATTCAGCTTAATTTG 1284
Db CCAGGAGCTGACTGTGGCTCAGATCATGAATCTCTTATTCGCAAAATTCAGCTTAATTTG 60307

QY 1285 AAGAAAGTAGGAAACCACTAGATCACTCAGTGAAGACCTAAATCCAATCCCTTTATGAT 1344

Db 60308 AAAAAAGTAGGAAAAACCACTAGACCACTCAGTATGACCTTAATCAATCCCTTATGAT 60367
QY 1345 TATACAGTGAAGTGAAGAAATAGATTTAAGGCCCTAGATCTGTATAGACAGAGTACCTAAT 1404
Db TATACAGTGAAGTGAAGAAATAGATTTAAGGCCCTAGATCTGTATAGATAGAGTACCTGAT 60427

QY 1405 GAACCTATGGACAGAGGTTTCATGACATTTGTACAGGAGACAGGATCGAGACCATCCCATG 1464
Db GAACCTATGGAAATGAGGTTTCGTCGACATTTGTACAGGAGACAGGATCAAGACCATCCCATG 60487

QY 1465 GAAAAAGAAATGCAAAAAACAAAAATGGCTGTCTGGGAGGCCCTTACAAATAGCTGTGAAA 1524
Db GAAGAGAAATGCAAAAAACAAAAATGCAAGTGGCTGTCTGGGAGGCCCTTACAAATAGCTGTGAAA 60547

QY 1525 AGAAGAGAAATGCAAAAAACAAAAAGAAAGATAAAGATTAAGATCTGAAATGCAAGATTC 1584
Db AGAAGAGAAATGCAAAAAACAAAAAGAAAGATTAAGATCTGAAATGCAAGATTC 60607

QY 1585 CAAA----- 1588
Db CAAAGATAGCAAGGAGAGATAAGAAAGCCCTTCTCAGCGATCACTGCAAAAGAAATAGAG 60667

QY 1589 ----- 1588
Db GAAAAACAAAGAAATGGGAAAGACTAGAGATCTCTTCAAGAAAAATTAGAGATACCAAGGGA 60727

QY 1589 ----- 1588
Db ATATTTTCATGCAAGATGGGTCGATTAAGGACAGAAATGGTATGGACCTAACAGAGCA 60787

QY 1589 ----- 1588
Db GAAGATATTAAGAAAGATCTGCAAGATATACAGAGAACTGTACAAAAAGATCTTTCAT 60847

QY 1589 ----- 1588
Db 1589 ----- 1588

Db 60848 GACCAGATAATACTCTGGTGTGATCACTGACCTAGAGCCAGACATCTCTGGACTGTGAA 60907

QY 1589 ----- 1588
Db 1589 ----- 1588

Db 60908 GTCAAGTGGCCCTTAGAAAGCATCACTACGAAACAAAGCTAGTGGAGGTGATGGAATCCA 60967

QY 1589 ----- 1588
Db 1589 ----- 1588

Db 60968 GTTGAGCTCTTCCAAATCCTGAAAGATGATGCTGTGAAAGTGTGCACTCAATATGCCAG 61027

QY 1589 ----- 1588
Db 1589 ----- 1588

Db 61028 CAAATTTGCAACTCAGCAGTGGCCACAGACTGGAAAGGTTAGTTTTCATTCCAATC 61087

QY 1589 ----- 1588
Db 1589 ----- 1588

Db 61088 CCAAGAAAGGCAATGCCAAAGAAATGCTCAAACTACTGCAAAATTTGCACTCATCTCACAC 61147

QY 1589 ----- 1588
Db 1589 ----- 1588

Db 61148 GCTAGTAAGTAATGCTCAAAATTTCTCAAGCCAGGCTTCAGCAATACCTTGAACCGGTGAA 61207

QY 1592 CTTCCAGTGTGTTCAAGCTGGTTTTAGAAAAAGTACAGAGAAACAGAGACCAAAATTCGCAAC 1651
Db CTTCCAGATGTTCAAGCTGGTTTTAGAAAAAGGAGAGAAACAGAGATCAAAATTCGCAAC 61267

QY 1652 ATCTCTGTATCATGGAAGAAAGCAAGAGTTCCAGAAAAACATCTATTTCTGCTTTATTT 1711
Db ATCCACTGGATCATGGAAGAAATGCAAGAGAGTTCCAGAAAAAGCATCTATTTCTGCTTTATTT 61327

QY 1712 GACTATGCAAAAGGCTTTGACTGTGGGGGTCAAAATAAACTGTGAAAAATTTCTGAAAAAGG 1771
Db GACTATGCTAAAGGCTTTGACTGTGGAATCAAAATAAACTGTGAAAAATTTCTGAAAAAGG 61387

QY 1772 ATGGGAATACAGAACCACTGACCTGACTCTTTGAAAAATTTGTATGAGGTGAGGTCAGAGCA 1831

```

Db 61388 ATGGGAATACAGACCATCTCTCTGCTCTTGAGAAACCTATATGCAGGTCAGGAAGCA 61447
QY 1832 ACAGTTAGAACTGGACATGGAACACACAGACTGGTTCCAAAGTAGGAAAAGAGATGTCAA 1891
Db 61448 AAAGTTAATAGCTGATATGGAACACACAGACTGGTTCCAAATAGGAAAAGAGTACATCAA 61507
QY 1892 GCCTGTATATTGTCTACCCGGCTTGTTAACTTCTATGCAGAG-ACATCATGAGAAACGCT 1950
Db 61508 GCCTGTATATTGTCTACCTGCTTATTAACTTATATGCAGAGTACATCATGAGAAACGCT 61567
QY 1951 GGGCTGGAAGACACAAAGCTGGATCAAGATTCAGATTCGGGGAGAAATAGCAATAAAGCTCTGATG 2070
Db 61568 GGGTTGGAAGAGACACAAAGCTGGATCAAGATTCAGATTCGGGGAGAAATATCAA-AAACTCAGA 61626
QY 2011 TATGCAGATGATACACACCTTATGCGAGAAAGTGAAGAGAACTAAAGAGCTCTTCTGATG 2070
Db 61627 TATGCAGATGATACACACCTTATGCGAGAAAGTGAAGAGAACTAAAGAGCTCTTCTGATG 61686
QY 2071 AAGTGAAGAGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAACGAAGATC 2130
Db 61687 AAAGTGAAGAGAGAGTGAAGAAAGTTGGCTTAAAGCTCAATATTCAGAAACCTAAGATC 61746
QY 2131 ATGCGATCTGCTCCATCACTTCATGGGAATAGATGGGAAACAGTGGAAACAGTGTCA 2190
Db 61747 ATGCGATCTGCTCCATCACTTCATGGGAATAGATGGGAAACAGTGGAAACAGTGTCA 61797
QY 2191 GACTTTATTTTGGGGGGCTCCAAATCACTGCAGATGTGACTGCAGCCATGAAATTA 2250
Db 61798 GACTTTATTTT--CTGGGCTCCAAATCACTGCAGATGTGACTGCAGCCATGAAATTA 61855
QY 2251 AAAGACACTTACTCTTGGGAAGAAAGTTA--ACCACTAGATAGATATTTGAAAGCA 2308
Db 61856 AAAGACGCTTACTCTTGGGAAGAAAGTTATGACCAACCTAGATAGCATATTTCAAAGCA 61915
QY 2309 GAGACATTACCTTGCACAAAGCCCATCTAGTCAAGGCTATGTTTTTCCAGTGTCTCA 2368
Db 61916 GAGACA-----TTGCCACAAAGTCCATCTAGTCAAGGCTATGTTTTTCCAGTGTCTCA 61970
QY 2369 TGTATGATGAGAGTTGGAATGTGAAGAAAGCTGAGCACTGAAGAAATGATGCTTTTG 2428
Db 61971 TGTATGATGAGAGTTGGAATGTGAAGAAAGCTGAGCGCTGAAGAAATGATGCTTTTG 62030
QY 2429 AACTGTGTGTGGAGAGACTCTTGAGAGTCCCTTGGCTGCAAGAGATCCAAACAGT 2488
Db 62031 AACTGTGTGTGGAGAGACTCTTGAGAGTCCCTTGGCTGCAAGAGATCCAAACAGT 62090
QY 2489 CCATTCTGAAGAGATCAGCCCTGGGATTTCTTTGGAAGAAATGATGCTAAAGCTGAAC 2548
Db 62091 CCATTCTGAAGAGATCAGCCCTGGGATTTCTTTGGAAGAAATGATGCTGAAGCTGAAC 62150
QY 2549 TCCAGTACTTTGGCCACCTGATCAGAGAGCTGACTCAGTGGAAAGACCCGATGCTGG 2608
Db 62151 TCCAGTACTTTGGCCACCTGATGTGAAGAGTTGACTCATTGGAAAGAGCTCTGATGCTGG 62210
QY 2609 GAGGATTTGGGGCAGGAGAGAGAGGACACAGAGATGAGATGCTGGATGGCATCA 2668
Db 62211 GAGGATTTGGGGCAGGAGAGAGAGGACACAGAGATGAGATGCTGGATGGCATCA 62270
QY 2669 CTGACTCGATGGAAGTGTGCTGGTGAATCTCTGAGTCTGGATGGTATGATGACAGGAGCCT 2728
Db 62271 CTGA--TCGATGGACATGATCTCAGTGAACCTCTGGAGTTGGTGTGATGACAGAGAGCCT 62329
QY 2729 GTCTGGGGCGGATTCATGGGCTCAAGAGTGTGGACAGCTGAGCAACTGACTGAAC 2788
Db 62330 GGCCTGCTGCCATTCATGGGGTCAAAAGAGTCCGACACGACTGAGGACTGAACTGAAC 62389
QY 2789 TGA 2791
Db 62390 TGA 62392

```

RESULT 42

AC109787/c

```

LOCUS AC109787 154577 bp DNA linear HTG 05-DEC-2002
DEFINITION Bos taurus clone RP42-136015, WORKING DRAFT SEQUENCE, 3 ordered
pieces.
ACCESSION AC109787
VERSION AC109787.2 GI:26050954
KEYWORDS HTG; HTGS PHASE2; HTGS_DRAFT.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 154577)
Benjamin, B., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,
Carlaga, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J.,
Kathig, P., Lee-Hin, S.-Q., Legaspi, R., Maduro, O.L., Maduro, V.B.,
Margaritis, E.H., Masello, C., Maskeri, B., McDowell, J.,
Parguirian, C., Pearson, R., Portnoy, M.E., Prasad, A.,
Reddi-Dugue, N., Schandler, K., Schueler, M.G., Sison, C.,
Stantridop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L.,
Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 154577)
Green, E.D.
Direct Submission
Submitted (07-FEB-2002) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 154577)
Green, E.D.
Direct Submission
Submitted (05-DEC-2002) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
On Dec 5, 2002 this sequence version replaced gi:18581925.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@nhgri.nih.gov
----- Project Information
Center project name: cjr
Center clone name: 136015

```

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

```

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 154307 bases at least Q40
Consensus quality: 154347 bases at least Q30
Consensus quality: 154358 bases at least Q20
Insert size: 121000; agarose-fp
Quality coverage: 12.74x in Q20 bases; agarose-fp
Quality coverage: 9.99x in Q20 bases; sum-of-contigs
-----

```

* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have

* provided by the submitter.

* This sequence will be replaced

* By the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 88559: contig of 8859 bp in length

* 88560 88559: gap of unknown length

* 88660 12656: contig of 37997 bp in length

* 12657 12656: gap of unknown length

* 126757 154577: contig of 27821 bp in length.

* Location/Qualifiers

1. .154577

/organism="Bos taurus"

/mol_type="genomic DNA"

/db_xref="taxon:9913"

/clone="RP42-136015"

/clone_lib="RP42"

1. .88559

/note="assembly_fragment"

clone_end:SP6

vector_side:left"

1. .14523

/note="clone overlaps with GenBank Accession Number

AC109792 clone RP42-475P24 (center project name cjm)"

57117. .154577

/note="clone overlaps with GenBank Accession Number

AC109798 clone RP42-71E11 (center project name cjq)"

88660. .12656

/note="assembly_fragment"

126757. 154577

/note="assembly_fragment"

clone_end:17

vector_side:right"

Query Match 52.2%; Score 1864.4; DB 1; Length 154577;

Best Local Similarity 76.0%; Pred. No. 0;

Matches 2288; Conservative 0; Mismatches 126; Indels 596; Gaps 5;

380 ATAGATCTGAAGAGAACTCCCGAGTCAAGTCAAGTGGTCCCATATGCTACTGGAGATC 439

87521 ATAGATCTGAAGAGAACTCCCGAGTCAAGTCAAGTGGTCCCATATGCTACTGGAGATC 87462

440 AGTGAGAAATAACTCCAGAAAGATGAAGATGGAGCCAAACCAAAAGAAATACCCAG 499

87461 AGTGAGAAATAACTCCAGAAAGATGAAGATGGAGCCAAACCAAAAGAAATACCCAG 87402

500 CTGTGGATGTGACTGGTGATATAAGCAAGTCCGATCTGTAAAGACAATATTTGCATAG 559

87401 CTGTGGATGTAACTGGTGATAGAGCAAGTCTGTATGCTGTAAAGACAATATTTGCATAG 87342

560 GAACCTGAATGTCAAGTCCATGATCAAGCAATTTGGAAGTGGTCAAAACAGAGATGG 619

87341 GAACCTGAATGTCAAGTCCATGATCAAGCAATTTGGAAGTGGTCAAAACAGAGATGG 87282

620 CAAGAGTGAATGTCAACATTTAGGAATTCAGCAACTAAATAGCAATGGATGGTGAAT 679

87281 CAAGAGTGAATGTCAACATTTAGGAATTCAGCAACTAAATAGCAATGGATGGTGAAT 87222

680 TTAAGTCAAGTCAATTTATTTACTTACTCTCGGCGCAGGATCCCTCAGAGAAATGGAG 739

87221 TTAAGTCAAGTCAATTTATTTACTTACTCTCGGCGCAGGATCCCTCAGAGAAATGGAG 87162

740 TAGCCATCATGTGCAACAAAGAGTCCGAATAGCAATGCAATGCAATGCAATGCAATGCA 799

87161 TGGCCATCAAGTCAACAAAGAGTCCGAATAGCAATGCAATGCAATGCAATGCAATGCA 87102

800 ACAGAAATCATCTGTGTTTTCAGGCAACCAATTCATATCATAGTATTCAGTCT 859

87101 ACAGAAATCATCTGTGTTTTCAGGCAACCAATTCATATCATAGTATTCAGTCT 87042

860 ATGCCCCCAACAGTAACTCTGAAGAGTGAAGTTGAACGCTCTATGAAGACCTTCAAG 919

87041 ATGCCCCCAACAGTAACTCTGAAGAGTGAAGTTGAACGCTCTATGAAGACCTTCAAG 86982

QY 920 ACCTTTTAGAATAAACAACCCCAAAAAGATGCTCTTCTCATTTATAGGGACTGGAATGCAA 979

Db 86981 ACCTTTTAGAATAAACAACCCCAAAAAGATGCTCTTCTCATTTATAGGGACTGGAATGCAA 86922

QY 980 AAGTAGGAAGCAAGAACAACCTGGAGTAAACAGGCAAAATTTGGCTTGGATACGGATG 1039

Db 86921 AAGTAGGAAGTCAAGAACAACCTGGAGTAAACAGGCAAAATTTGGCTTGGATACGGATG 86862

QY 1040 AAGCAGGCGCAAGACTAATAGAGTTTTCGCCAAGAAATGCACTGGTCTATAGCAACACCC 1099

Db 86861 AAGCAGGCGCAAGACTAATAGAGTTTTCGCCAAGAAATGCACTGGTCTATAGCAACACCC 86802

QY 1100 TCTTCCAAACAACAAGAGAGACTCTACACATGGACATCAACAGATGGTCAACACCGAA 1159

Db 86801 TCTTCCAAACAACAAGAGAGACTCTACACATGGACATCAACAGATGGTCAACACCGAA 86742

QY 1160 ATCAGATTGATTATATTTCTTTCAGCCCAAGATGGAGAGCTCTATACAGTCAGCAAAA 1219

Db 86741 ATCAGATTGATTATATTTCTTTCAGCCCAAGATGGAGAGCTCTATACAGTCAGCAAAA 86682

QY 1220 CAAGACCCAGGAGCTTACTGTGCTCAGATCATGAATCTCTTATTGCCAAATTCAGACTTA 1279

Db 86681 CAAGACCCAGGAGCTTACTGTGCTCAGATCATGAATCTCTTATTGCCAAATTCAGACTTA 86622

QY 1280 AATTGAAGAAAGTAGGGAACCACTAGATCACTCAGGTAGACCTTAATCCATCCCTT 1339

Db 86621 AATTGAAGAAAGTAGGGAACCACTAGATCACTCAGGTAGACCTTAATCCATCCCTT 86562

QY 1340 ATGATTATACAGTGAAGTGAAGAAATAGATTTAAGGGCTTAGATCTGTAGACAGATAC 1399

Db 86561 ATGATTATACAGTGAAGTGAAGAAATAGATTTAAGGGCTTAGATCTGTAGATAGATGC 86502

QY 1400 CTAATGAACATATGGACAGAGTTTCAATGATTTGACAGGACACAGGATCGAGACCATCC 1459

Db 86501 CTGATGAGCTATGGATGAGTTCTGTGATTTGTACAGGACACAGGATCGAGACCATCC 86442

QY 1460 CCATGGAAGAAAGTGAAGAAAGCAAAATGGCTCTCTGGGAGGCTTCAAAATAGCTG 1519

Db 86441 CCATG - AAGAAATGCAAAAGCAAAATGGCTCTCTGGGAGGCTTCAAAATAGCTG 86383

QY 1520 TGAAGAGAGAGAGTGAAGCAAGGAAAGAAAGAAAGATATAAGCATCTGAATGAG 1579

Db 86382 TGAAGAGAGAGAGTGAAGCAAGGAAAGGAGAAAGAAAGATATAAGCATCTGAATGAG 86323

QY 1580 AGTTCCAA - 1588

Db 86322 AGTTCCAAAGATAGCAAGAGATAGAAAGCCCTTCTCTAGTGATCAATGCAAGAA 86263

QY 1589 - 1588

Db 86262 TAGAGGAAACAACAGATGGGAAAGACTAGAGATCTCTTCAAGAAATTCAGAGATACCA 86203

QY 1589 - 1588

Db 86202 AAGGAACATTTTCATGCAAGATGGGCTCAATAAGGACAGAAATGGTATGGACCTAACAG 86143

QY 1589 - 1588

Db 86142 AAGCAGAGATATTAAAGAGAGATGGCAAGATACACAGAGAACTGTACAAAAAAGATC 86083

QY 1589 - 1588

Db 86082 TTCAATGACCCAGATATCACGATGGTGTGATCACTGACCTAGAGCCAGACATCTGGACT 86023

QY 1589 - 1588

Db 86022 GTGAGTCAAGTGGGCTTAGAATCACTATGCAAAAGCTAGTGGAGTGAATTC 85963

QY 1589 - 1588

Db 85962 AGTTGAGCTATTTCAAATCCTAAAGATGATGCTGTGAAAGTGTGCTCAATATGCCA 85903

QY 1589 - 1588

Db 85902 GCATAATTTGGAATACTCAGCAGTGGCCACAGGACTGGAAAAGGTGAGTTTCAATTCAT 85843
Qy 1589 ----- 1588
Db 85942 CCCAAAGAAAGCAATGTCAAAGAAATGCTCAAACTACCGCACAAATGCACTCATCTCAC 85783
Qy 1589 -----GA 1590
Db 85782 CGCTAGTAAGTAATGTTAAATTTCCAAAGCAGCGCTTCAGCAATCTTTGAACCATGA 85723
Qy 1591 ACTTCCAGTTGTTTCAAGCTGTTTTAGAAAAGTCAGAGGAACACAGAGACCAAAATTCGCCAA 1650
Db 85722 ACTTCCAGTTGTTTGAAGCTGTTTTAGAAAAGGCAGACGACAGATCAAAATTCGCCAA 85663
Qy 1651 CATCCTCTGTATCATGGAATAAGCAAGAGAGTTCCAGAAAACATCTATTTCTGCTTTAT 1710
Db 85662 CATATGCTGGATCATGGAATAAGCAAGAGAGTTCCAGAAAACATCTATTTCTGCTTTAT 85603
Qy 1711 TGACTATGCAAAAGCCTTTGACTGTGGGGGTCACAATAAATCTGTGGAATAATCTGAAAGG 1770
Db 85602 TGACTATGCAAAAGCCTTTGACTGTGGGTGATCATCAATAAATCTGTGGAATAATCTGAAAG 85543
Qy 1771 GATGGGAATACCAAGCAACCTGACTGACTCTTGAAAATTTGTATGCAAGTCAAGAAC 1830
Db 85542 GATGGGAATACCTGACCTCTGACTGCTCTTGAGAAATCTGTATGCAGATCAGGAAGC 85483
Qy 1831 AACAGTTAGACTGACATGGAACAACAGACTGTTCCAAAGTAGGAAAAGGATGATGCA 1890
Db 85482 AACATTAGAACTGGACATGGAACAGCAGACTGATTCCTAATAGAAAAGGATGATGCA 85423
Qy 1891 AGGCTGTATATGTCACCCGCTTTGTTAACTTCTATGCAGAG-ACATCATGAAACGC 1949
Db 85422 AGGCTGTATATGTCACCCCTCTTATTTAACTTCTATGCAGAGTACATCATGAAACGC 85363
Qy 1950 TGGCTGGAAGAACAACAAGTGAATCAAGATTGCCGGGAGAAATACCAATACCTCAG 2009
Db 85362 TGGCTGGAAGAACAACAAGTGAATCAAGATTGCCGGGAGAAATACCAATACCTCAG 85303
Qy 2010 ATATGCAATATACCAACCTTTATGGGAGAAAGTGAAGAGAACTAAAGAGCTCTTGAT 2069
Db 85302 ATATGCAATATACCAACCTTTATGGGAGAAAGTGAAGAGAACTAAAGAGCTCTTGAT 85243
Qy 2070 GAAGGTGAAAGAGAGAGTGAAGAAAGTGGTTTAAAGCTCAACATTGAAAGCAAGAT 2129
Db 85242 GAAGGTGAAAGAGAGAGTGAAGAAAGTGGTTTAAAGCTCAACATTGAAAGCAAGAT 85183
Qy 2130 CATGCACTCTGCTCCCATCACTTCATGGAAATAGATGGGAAACAGTGGAAACAGTGTC 2189
Db 85182 CATGCACTCTGCTCCCATCACTTCATGGCAATAGATGGGAAACAGTGGAAACAGTGTC 85123
Qy 2190 AGACTTTATTTTTGGGGGGCTCCAAAATCACTGCAGATGGTGAAGTCCAGCCATGAAT 2249
Db 85122 AGACT---TATTTGGGGGGCTCCAAAATCACTGCAGATGGTGAAGTCCAGCCATGAAT 85066
Qy 2250 AAAGACACTTACTCTTTGGAAGAAAGTTA--ACCACTAGATAGATGGAAGTCCAGTGAAGC 2307
Db 85065 AAAGACACTTACTCTTTGGAAGAAAGTTATGACCAACCTAGATAGATGGAAGTCCAGTGAAGC 85006
Qy 2308 AGAGACATTACTTCCCAACAAAGCCCATCTAGTCAAGGCTATGTTTTTCCAGTGGTC 2367
Db 85005 AGAGACATTATTTGCCAACAAAGGTCTGCTAGTCAAGGCTATGTTTTTCCAGTGGTC 84946
Qy 2368 ATGATGATGATGAGAGTTGACTGTGAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2427
Db 84945 ATGATGATGATGAGAGTTGACTGTGAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 84886
Qy 2428 GAATGCTGCTGTGAGAGAGCTTTGAGAGTCCCTTGGACTGCAAGAGAGATCAACACAG 2487
Db 84885 GAATGCTGCTGTGAGAGAGCTTTGAGAGTCCCTTGGACTGCAAGAGAGATCAACACAG 84826
Qy 2488 TCCATTCTGAAGGAGATCAGCCCTGGGATTTCTTTGGAAGGAATGATGCTAAAGCTGAAA 2547

Db 84825 TCCATTCTGAAGGAGATCAGCCCTGGGATTTCTTTGGAAGGAATGATGCTAAAGCTGAAA 84766
Qy 2548 CTCAGTACTTTGGGCCACCTGATCAGAAAGCTGACTCACTGGAAGACCTGATGCTG 2607
Db 84765 CTGTAGTACTTTGGGCCACCTTATGCGAAAGTTGACTCATTTGGAAGAACTCTGATGCTG 84706
Qy 2608 GGAGGGATTTGGGGGAGAGAGGAGGAGGACACAGAGGATGAGATGGCTGATGGCATC 2667
Db 84705 GGAGGGATTTGGAGCAGAGAGGAGGAGGATGACAGAGGATGAGATGCTGGATGGCATC 84646
Qy 2668 ACTGACTGATGGAGCTGAGTCTGGTGAACTTCCTGGAGTTGGTGTATGACAGAGGAGGCC 2727
Db 84645 ACTGACTGATGGCGCTGAGTCTGAGTGAATCTCTGGAGTTGGTGTATGACAGAGGAGGCC 84586
Qy 2728 TGTCTTGGCGGATTCATGCGGTCAAAAGATTTGGACACGACTGAGCAACTGAACTGAA 2787
Db 84585 TGGCTGTCTGCAATTCATGCGGTCCAAAGAGTCCGACACGACTGAGTGAATGAACTGAA 84526
Qy 2788 CTGAACCTGTA 2797
Db 84525 CTGAACCGAA 84516

RESULT 43
AC109798/c 201118 bp DNA linear HTG 05-DEC-2002
LOCUS Bos taurus clone RP42-71E11, WORKING DRAFT SEQUENCE, 5 ordered
DEFINITION pieces.

AC109798
AC109798.2 GI:26050957
HTG; HTGS_PHASE2; HTGS_DRAFT.

KEYWORDS
SOURCE Bos taurus (cow)

ORGANISM
Bos taurus

REFERENCE
AUTHORS
Bovidae; Bovinae; Bos
1 (bases 1 to 201118)

Akheri, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Carliaga, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J.,
Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E.,
Laric, P., Lee-Lin, S.-O., Legaspi, R., Maduro, O.L., Maduro, V.B.,
Margulies, E.H., Masiello, C., Maskeri, B., McDowell, J.,
Pagnir-Gane, C., Pearson, R., Portnoy, M.E., Prasad, A.,
Reddy-Dugue, N., Schandler, K., Schreier, M.G., Sison, C.,
Stathropop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L.,
Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.

NISC Comparative Sequencing Initiative

Unpublished

2 (bases 1 to 201118)

Green, E.D.

Direct Submission

Submitted (07-FEB-2002) NIH Intramural Sequencing Center, 8717

Groveton Circle, Gaithersburg, MD 20877, USA

3 (bases 1 to 201118)

Green, E.D.

Direct Submission

Submitted (05-DEC-2002) NIH Intramural Sequencing Center, 8717

Groveton Circle, Gaithersburg, MD 20877, USA

On Dec 5, 2002 this sequence version replaced gi:18582035.

----- Genome Center

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: <http://www.nisc.nih.gov>

Contact: nisc_zoo@nhgri.nih.gov

----- Project Information

Center project name: cjq

Center clone name: 071E11

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair

data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross missassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics -----

Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 199764 bases at least Q40
 Consensus quality: 200000 bases at least Q30
 Consensus quality: 200271 bases at least Q20
 Insert size: 157000; agarose-fp
 Insert size: 200718; sum-of-contigs
 Quality coverage: 12.42x in Q20 bases; agarose-fp
 Quality coverage: 9.72x in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 5 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submittor.

* This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

* 1 70538: contig of 70538 bp in length
 * 70539 70638: gap of unknown length
 * 70639 77542: contig of 6904 bp in length
 * 77543 77642: gap of unknown length
 * 77643 133416: contig of 55774 bp in length
 * 133417 133516: gap of unknown length
 * 133517 149909: contig of 16393 bp in length
 * 149910 150009: gap of unknown length
 * 150010 201118: contig of 51109 bp in length.

FEATURES

source

1. .201118
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9913"
 /clone="RP42-71E11"
 /clone_lib="RP42"

misc_feature

1. .98740
 /note="clone overlaps with GenBank Accession Number
 AC109787 clone RP42-136015 (center project name cjr)"

misc_feature

1. .70538
 /note="assembly_fragment
 clone_end:SP6
 vector_side:left"

misc_feature

70639. .77542
 /note="assembly_fragment"
 77643. .133416

misc_feature

133517. .149909
 /note="assembly_fragment"

misc_feature

150010. .201118
 /note="assembly_fragment
 clone_end:T7
 vector_side:right"

misc_feature

153492. .201118
 /note="clone overlaps with GenBank Accession Number
 AC109797 clone RP42-58E2 (center project name cjp)"

Query Match 52.2%; Score 1864.4; DB 1; Length 201118;

Best Local Similarity 76.0%; Pred.No. 0;

Matches 2288; Conservative 0; Mismatches 126; Indels 596; Gaps 5;

Qy

380 ATAGGACTCAAGAGAACTCCCGAGGTGAGTGGCCCATGCTACTGGAGATC 439

Db

30436 ATAGGACTCAAGAGAACTCCCGAGGTGAGTGGCCCATGCTACTGGAGATC 30377

Qy	440	AGTGGAGAAATAACTCCAGAAAAGAAATGAAGAGATGAGCCAAAGCAAAAAGAAATACCCAG	499
Db	30376	AGTGGAGAAATAACTCCAGAAAAGAAATGAAGAGATGAGCCAAAGCAAAAAGAAATACCCAG	30317
Qy	500	CTGTGGATGTGACTGGTGTATATAAGCAAGGTCCGATGCTGTAAAGAGCAATATTGTCATAG	559
Db	30316	CTGTGGATGTGACTGGTGTATATAAGCAAGGTCTGTATGCTGTAAAGAGCAATATTGTCATAG	30257
Qy	560	GAACCTGGATGTGAGTCCATGAATCAGGCAAAATGGAGTGGTCAACAGAGATGG	619
Db	30256	GAACCTGGATGTGAGTCCATGAATCAGGCAAAATGGAGTGGTCAACAGAGATGG	30197
Qy	620	CAAGAGTGAATGTCAACATTTCTAGGAATCAGGAACTAAATAGGACTGGATGGTGAAT	679
Db	30196	CAAGAGTGAATGTCAACATTTCTAGGAATCAGGAACTAAATAGGACTGGATGGTGAAT	30137
Qy	680	TTAACTCAGATGACCATTTATATCTACTACTTGGGGCAGGAATCCCTCAGAGAAATGGAG	739
Db	30136	TTAACTCAGATGACCATTTATATCTACTACTTGGGGCAGGAATCCCTCAGAGAAATGGAG	30077
Qy	740	TAGCCATCATGTCAACAAAGAGTCCGAAATGCAAGTACTTGGATGCACTCTCAAAAACG	799
Db	30076	TGCGCATCAGGTCAACAAAGAGTCCGAAATGCAAGTACTTGGATGCACTCTCAAAAACG	30017
Qy	800	ACAGAATGATCTCTGTGTTTCCAAAGGCAACCAATCAATATACAGTATATCAAGTCT	859
Db	30016	ACAGAATGATCTCTGTGTTTCCAAAGGCAACCAATCAATATACAGTATATCAAGTCT	29957
Qy	860	ATGCCCCACCAAGTAACTGTAAGAGCTGAAGTTGAACGGTCCCTATGAAGACCTTACAAAG	919
Db	29956	ATGCCCCACCAAGTAACTGTAAGAGCTGAAGTTGAACGGTCCCTATGAAGACCTTACAAAG	29897
Qy	920	ACCTTTTGAAGTAACTAAACACCCCAAAAAGAGTCTCTTCTCATATAGGGGACTGGAATGAA	979
Db	29896	ACCTTTTGAAGTAACTAAACACCCCAAAAAGAGTCTCTTCTCATATAGGGGACTGGAATGAA	29837
Qy	980	AGTAGGAAGCAAGAACACCTGGAGTAAACAGGCAAAATTTGGCTTTGGATACGGATG	1039
Db	29836	AGTAGGAAGTCAAGAAACAGCTGGAGTAAACAGGCAAAATTTGGCTTTGGATACGGATG	29777
Qy	1040	AAGCAGGGCAAGAGACTAATAGAGTTTTCAGAGAAATGCACTGGTCTATAGCAACACCC	1099
Db	29776	AAGCAGGGCAAGAGACTAATAGAGTTTTCAGAGAAATGCACTGGTCTATAGCAACACCC	29717
Qy	1100	TCTTCCAAACACAAAGAGAGACTCTTACATGAGACATCACAGATGGTCAACACCGAA	1159
Db	29716	TCTTCCAAACACAAAGAGAGACTCTTACATGAGACATCACAGATGGTCAACACCGAA	29657
Qy	1160	ATCAGATTGATTATATCTTTTCAGCCAAAGATGGAGAGCTCTATACAGTCAGCAAAA	1219
Db	29656	ATCAGATTGATTATATCTTTTCAGCCAAAGATGGAGAGCTCTATACAGTCAGCAAAA	29597
Qy	1220	CAAGACAGGAGCTTACTGTGGCTCAGATCATGAACTCCTTATTTGCCAAATTCAGACTTA	1279
Db	29596	CAAGACAGGAGCTTACTGTGGCTCAGATCATGAACTCCTTATTTGCCAAATTCAGACTTA	29537
Qy	1280	AAATGAAGAAAGTAGGAAACCACTAGATCACTCAGGTAGAGCTTAATCAATCCCTT	1339
Db	29536	AAATGAAGAAAGTAGGAAACCACTAGATCACTCAGGTAGAGCTTAATCAATCCCTT	29477
Qy	1340	ATGATTATACAGTGGAGATGAGAAATAGATTTTAAGGGCTTAGATCTGTATAGACAGATAC	1399
Db	29476	ATGATTATACAGTGGAGATGAGAAATAGATTTTAAGGGCTTAGATCTGTATAGATAGATGC	29417
Qy	1400	CTAATGAATATGACAGAGGTTCATGACATTTACAGAGACAGAGGATCCAGACCATCC	1459
Db	29416	CTGATGACTATGGAATGAGGTTCGTGACATTTGTACAGGACAGAGGATCCAGACCATCC	29357
Qy	1460	CCATGGAAAAAGAAATGCAAAAAGCAAAATGGCTGTCTGGGGAGGCCCTTACAAATAGCTG	1519
Db	29356	CCATGG-AAAAGAAATGCAAAAAGCAAAATGGCTGTCTGGGGAGGCCCTTACAAATAGCTG	29298

```
QY 1520 TGAAGAAGAGAGTGAAGAAAGCAAGAAAGAAAGAAAGAAAGCAATCTGAATGCAG 1579
Db 29297 TGAAGAAGAGAGAGCAAGAAAGCAAGAAAGAAAGAAAGAAAGTATAAACAATCTGAATGCAG 29238
QY 1580 AGTTCCAAA----- 1588
Db 29237 AGTTCCAAAAGATAGCAAGAGAGATAGAAAGCCTTCTCAGTGATCAATGCAAGAA 29178
QY 1589 ----- 1588
Db 29177 TAGAGAAACAACAAGATGGAAGAGACTAGAGATCTTCAAGAAATCAGAGATACCA 29118
QY 1589 ----- 1588
Db 29117 AAGGAACATTTATGCAAGATGGCTCAATAAAGGACAGAAATGATGATGATCAACAG 29058
QY 1589 ----- 1588
Db 29057 AAGCAGAGATATTAGNAGAGATGGCAAGATACACAGAGACTGTACAAAAGATC 28998
QY 1589 ----- 1588
Db 28997 TTTATGACCCAGATAATCAGATGGTGTGATCACTGACCTAGAGCCAGACATCCTGGACT 28938
QY 1589 ----- 1588
Db 28937 GTGAAGTCAAGTGGCCCTTAGAATCACTATGGACAAAGCTAGTGGAGGTGATGGAATTC 28878
QY 1589 ----- 1588
Db 28877 AGTTGAGCTATTTCAAATCCTAAAAGATGATGTGTGAAGTGTGCACTCAATATGCCA 28818
QY 1589 ----- 1588
Db 28817 GCAAAATTTGAAACTCAGCAGTGGCCACAGGACTGGAAAGTCAAGTTTCATTCOAAT 28758
QY 1589 ----- 1588
Db 28757 CCCAAGAGAGCAATGTCAAAGATGCTCAAACCTACGGCACAATTGCACTCATCTACA 28698
QY 1589 -----GA 1590
Db 28697 CGCTAGTAAGGTAATGTTAAATTTCTCAAGCAGGCTTTCAGCAATACCTTGAACCATGA 28638
QY 1591 ACTTCCAGTGTCTCAAGCTGTTTTAGAAAAGTCAGAGGAAACAGACACCAAAATTCGCCAA 1650
Db 28637 ACTTCCAGTGTGAGAGTGTGTTTTAGAAAAGGACGACGACCAAGATCAAAATTCGCCAA 28578
QY 1651 CATCTCTGTATCATGGAAGAAAGCAAGAGATTTCCAGAAAACATCTATTTCTGCTTTAT 1710
Db 28577 CATATGCTGGATCATGGAAGAAAGCAAGAGATTTCCAGAAAACATCTATTTCTGCTTTAT 28518
QY 1711 TGACTATGCAAAAGCCTTTGACTGTGGGGTCACAATATAACTGTGGAATTTCTGAAGG 1770
Db 28517 TGACTATGCAAAAGCCTTTGACTGTGGGGTCACAATATAACTGTGGAATTTCTGAAGG 28458
QY 1771 GATGGGAATACAGACACCTGACTGACTCTTGAATAATTTGATGCAAGTCAAGGAGC 1830
Db 28457 GATGGGAATACCTGACCTCTGACTGCTCTTGAATAATTTGATGCAAGTCAAGGAGC 28398
QY 1831 AACAGTTAGAACTGGAACATGGAACAACAGACTGTTCCAGTAGGAAAAGAGATGTCA 1890
Db 28397 AACATTAGAACTGGAACATGGAACAACAGACTGTTCCAAATAGGAAAAGAGTACGTCA 28338
QY 1891 AGGTGTATATTGTCACCCGCTTTGTTAACTTCTATGTCAGAG-ACATCATGGAAGACGC 1949
Db 28337 AGGTGTATATTGTCACCCCTCTTATTAACTTCTATGTCAGAGTACATATGGAAGACGC 28278
QY 1950 TGGCTGGAAGAACCAAGCTGGAATCAAGATTTGCCGGGAGAAATAGCAATAACCTCAG 2009
Db 28277 TGGCTGGAAGAACCAAGCTGGAATCAAGATTTGCCGGGAGAAATATCAATAACCTCAG 28218
QY 2010 ATATGCAGATATACCACCTTTATGGCAGAAAGTGAAGGAACTAAAAGCCTCTTGAT 2069
```

```
Db 28217 ATATGCAGATGACACCCCTTTATGCGAAAAGTGAAGAGAACTAAAAGCCTCTTGAT 28158
QY 2070 GAAGGTGAAGAGGAGAGTGAAGAAAGTGGCTTAAAGCTCAACATTCAGAAACGAAGAT 2129
Db 28157 GAAGGTGAAGAGGAGAGTGAAGAAATGTTGGCTTAAACTCAACATTCGGAACGAAGAT 28098
QY 2130 CATGCGATCTGTCCTCCATCATTGCGAAATAGATGGGAAACAGTGGAAACAGTGTGTC 2189
Db 28097 CATGCGATCTGTCCTCCATCATTGCGAAATAGATGGGAAACAGTGGAAACAGTGTGTC 28038
QY 2190 AGACITTTATTTTGGGGGCTCCAAATCACTGCGAGATGGTGAATGCGAGCCATGAAT 2249
Db 28037 AGACT---TATTTGGGGGCTCCAAATCACTGCGAGATGGTGAATGCGAGCCATGAAT 27981
QY 2250 AAAAGACATTTACTCTCTGGAAGAAAGTTA--ACCACCTAGATAGCATATTGAAAAGC 2307
Db 27980 AAAAGACCTTACTCTCTGGAAGAAAGTTATGACCACTAGATAGCATATTGAAAAGC 27921
QY 2308 AGAGACATTTACTCTGCGCAACAAAGCCCATCTAGTCAAGGCTATGGTTTTCCAGTGTGTC 2367
Db 27920 AGAGACATTTATTTGCCAAACAAAGGTCTGTCTAGTCAAGGCTATGGTTTTCCAGTGTGTC 27861
QY 2368 ATGTATGATGTGAGAGTTGGACTGTGAAGAAAGCTGAGCACTGAAGAAATTCATGCTTTT 2427
Db 27860 ATGTATGATGTGAGAGTGGGACTGTGAAGAAAGCTGAGTCTGAAGAAATTCATGCTTTT 27801
QY 2428 GAAGTGTGTGTGAGAGAGACTCTTGAGAGTCCCTTGACTGCAAGGAGATCCAAACAG 2487
Db 27800 GAAGTGTGTGTGAGAGAGACTCTTGAGAGTCCCTTGACTGCAAGGAGATCCAAACAG 27741
QY 2488 TCCATTTCAAGAGAGATCAGCCCTGGATTTCTTTGGAGGAAATGATGCTAAAGCTGAA 2547
Db 27740 TCCATTTCAAGAGAGATCAGCCCTGGATTTCTTTGGAGGAAATGATGCTAAAGCTGAA 27681
QY 2548 CTCCAGTACTTTGGCCACCTGATCAGAGAGCTGACTCACTGCAAGAAAGCCCTGATGCTG 2607
Db 27680 CTGCTAGTACTTTGGCCACCTGATCAGAGAGTGTGACTCATTTGCAAGAGCTCTGATGCTG 27621
QY 2608 GGAGGAAATGGGGCAGAGGAGAGAGGACACAGAGATGAGATGGCTGATGGCATC 2667
Db 27620 GGAGGAAATGGGGCAGAGGAGAGAGGAGATGACAGAGATGAGATGCTGATGGCATC 27561
QY 2668 ACTGACTCCATGAGCTGAGTCTGGTGAATCTCTGGAGTCTGGTGTGATGACAGAGGAGC 2727
Db 27560 ACTGACTCCATGAGCTGAGTCTGGTGAATCTCTGGAGTCTGGTGTGATGACAGAGGAGC 27501
QY 2728 TGTCTGCGCGATTCATGGGTCAAAAGATTGCAACAGATGGAACAGCTGAGCAACTGAA 2787
Db 27500 TGGCGTGTGCAATTCATGGGTGCAAGAGTGGACACGACTGAGTGAATGAACTGAA 27441
QY 2788 CTGCACTGTA 2797
Db 27440 CTGCACTGTA 27431
```

RESULT 44

```
AC137653
LOCUS Bos taurus clone RP42-506N12, WORKING DRAFT SEQUENCE, 6 ordered
DEFINITION pieces.
ACCESSION AC137653
VERSION AC137653.2 GI:28261477
KEYWORDS HTG; HTGS PHASE2; HTGS_DRAFT.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
            Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 121609)
AUTHORS Akhter N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,
            Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
            Cariaga, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J.,
```


Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E.,
 Latic, P., Lee-Lin, S.-Q., Legaspi, R., Maduro, Q.L., Maduro, V.B.,
 Margulies, E.H., Masiello, C., Maskeri, B., McDowell, J.,
 Paquirigan, C., Pearson, R., Portnoy, M.E., Prasad, A.,
 Reddi-Dugue, N., Schandler, K., Schueler, M.G., Sison, C.,
 Startipop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L.,
 Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished
 NISC Comparative Sequencing Initiative
 2 (bases 1 to 121609)

Green, E.D.

Direct Submission

Submitted (27-NOV-2002) NIH Intramural Sequencing Center, 8717

Groveomont Circle, Gaithersburg, MD 20877, USA

3 (bases 1 to 121609)

Green, E.D.

Direct Submission

Submitted (06-FEB-2003) NIH Intramural Sequencing Center, 8717

Groveomont Circle, Gaithersburg, MD 20877, USA

On Feb 6, 2003 this sequence version replaced gi:25699809.

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: <http://www.nisc.nih.gov>

Contact: nisc.zoo@nhgri.nih.gov

----- Project Information

Center project name: crv

Center clone name: 506N12

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 119159 bases at least Q40
 Consensus quality: 120104 bases at least Q30
 Consensus quality: 120787 bases at least Q20
 Insert size: 113000; agarose-fp
 Insert size: 121109; sum-of-contigs
 Quality coverage: 10.42x in Q20 bases; agarose-fp
 Quality coverage: 9.72x in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently consists of 6 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

1 18392: contig of 18392 bp in length
 18393 18492: gap of unknown length
 18493 25477: contig of 6985 bp in length
 25478 25577: gap of unknown length
 25578 46406: contig of 20829 bp in length
 46407 46506: gap of unknown length
 46507 49074: contig of 2568 bp in length
 49075 49175: gap of unknown length
 49175 89136: contig of 39962 bp in length
 89137 89236: gap of unknown length
 89237 121609: contig of 32373 bp in length.

FEATURES		Location/Qualifiers
source	1..121609	/organism="Bos taurus"
		/mol_type="genomic DNA"
		/db_xref="taxon:9913"
		/clone="RP42-506N12"
		/clone_lib="RP42"
misc_feature	1..18392	/note="assembly_fragment"
	clone_end:T7	vector_side:left"
misc_feature	1..13224	/note="clone overlaps with GenBank Accession Number AC137655 clone RP42-542P14 (center project name cru)"
	18493..25477	/note="assembly_fragment"
misc_feature	25578..46406	/note="assembly_fragment"
	46507..49074	/note="assembly_fragment"
misc_feature	49175..89136	/note="assembly_fragment"
	89237..121609	/note="assembly_fragment"
misc_feature	clone_end:SP6	vector_side:right"
Query Match 52.1%; Score 1862.6; DB 1; Length 121609;		
Best Local Similarity 74.8%; Pred. No. 0; Mismatches 194; Indels 600; Gaps 8;		
Matches 2359; Conservative 0; Gaps 0; Indels 600; Gaps 8;		
QY	249	GTGGGGCCAAACCAAGATGGCAGGTCTATGGTGAGAGATCTGACAGAAATGGTCCACTG 308
Db	68527	GTGGGGCCAAACCAAGATGGCAGGTCTATGGTGAGAGATCTGACAGAAATGGTCCACTG 68586
QY	309	GAGAAGGGAAT-GCAACCACTTCAGTATCTTGTGAGACCCCATGACAGATGA 367
Db	68587	GAGAAGGGAATGAGCAATCTTCTGTGAGACCCCATGACAGATGA 68646
QY	368	AAAGGCAAAATGATAGGATCTGAAAGAGGAATCTCCAGGTCAGTAGTGCCTCATATG 427
Db	68647	AAAGGCAAAATGATAGGATCTGAAAGAGGAATCTCCAGGTCAGTAGTGCCTCATATG 68706
QY	428	CTACTGGAGATCAGTGGAGAAATACTCCAGAAAGATGAAGATGAGCCAAAGCAAA 487
Db	68707	CTACTGGAGATCAGTGGAGAAATACTCCAGAAAGATGAAGATGAGCCAAAG-TAA 68765
QY	488	AAAGATACCCAGCTGTGGATGACTGTGGTATATAGCAAGGTCCGATGCTGTAAGAGC 547
Db	68766	AAAGATACCCAGCTGTGGATGACTGTGGTATATAGCAAGGTCCGATGCTGTAAGAGC 68825
QY	548	AAATATTCATAGGAACCTGGAAATGTCCAGTCCATCAATCAAGGCAAAATGGAGTGTCA 607
Db	68826	AAATATTCATAGGAACCTGGAAATGTCCAGTCCATCAATCAAGGCTAACTGCAAGTGTCA 68885
QY	608	ACAAGAGATGCAAGAGTATGTCACATCTCTAGGATCAGCAACTAAATGAGTCACTG 667
Db	68886	ACAAGAGATGCAAGAGTATGTCACATCTCTAGGATCAGCAACTAAATGAGTCACTG 68945
QY	668	GAATGGGTGAATTTAACTCAGATGACCATTTATCTACTCTCGGGCAGGAATCCCTCA 727
Db	68946	GAATGGGTGAATTTAACTCAGATGACCATTTATCTACTCTCGGGCAGGAATCCCTCA 69005
QY	728	GAAGAAATGAGTACCATCATGTCACAAAGAGTCCGAATCAGTACTTGTGATGCA 787
Db	69006	GAAGAAATGAGTACCATCATGTCACAAAGAGTCCGAATCAGTACTTGTGATGTA 69065
QY	788	GTCTCAAAAACGACAGATGATCTCTGTTTGTTCCTCAAGGCAAAACCAATTAATCAG 847
Db	69066	ATCTCAAAATGACAGATGATC-----GTGTTTCTAAGGCAAAACCAATTAATCAG 69121
QY	848	TAATCCAGTCTATGCCCCCAACAGTATGCTGAAGAGCTGAAGTGAACGGTCTATG 907

Db 69122 TAAATCAAGTCTATGCCCCAACCAAGTAATGCTGAAGAAGCTGAAGTTGAAACAGTTCTATG 69181
Qy 908 AGAGCTTACAGACCTTTTATAGAACTAACAACCCAAAAAAGATGTCCTCTCTATTATAGGG 967
Db 69182 AAGAGCTTACAGACCTTTTATAGAACTAACAACCCAAAAAAGATGTCCTTTTCATTATAAGGG 69241
Qy 968 ACTGGAATGCAAAAGTAGGAAGCAAGAAACACCTGGAGTACAGGCAAAATTTGGCCTTG 1027
Db 69242 ACTGGAATGCAAAAGTAGGAAGTCAAGAAACACCTGGGGTACAGGCAAAATTTGGCAATTG 69301
Qy 1028 GAATACGGAATGAAGCAGCGCAAGAACTAATAGAGTTTGGCAAGAAATGCCTGGTCA 1087
Db 69302 GAACACGGAATGAAGCAGCGCAAGTCTAATAGACTTTTGGCAAGAAACGCCTGGTCA 69361
Qy 1088 TAGCAAAACACCTCTTCCAAACACACAGAGAGAGTCTTACATCGACATCAACAGATG 1147
Db 69362 TAGCAAAACACCTCTTCCAAACACACAGAGAGAGTCTTACATCGACATCAACAGATG 69421
Qy 1148 GTCAACACCGAAATCAGATTGATTATATCTTTTGCAGGCAAGATGGGAGAGTCTATAC 1207
Db 69422 GTCAACACCGAAATCAGATTGATTCTATTCTTTGCAGGCAAGATGGGAGAGTCTATAC 69481
Qy 1208 AGTCAGCAAAACACAGACAGAGCTTACTGTGGCTCAGATCATGAATCTCTTATGCA 1267
Db 69482 AGTTAGCAAAACACAGACAGAGCTGACTGTGGCTCAGATCATGAATCTCTTATGCA 69541
Qy 1268 AATTCAGACTTAAATTTGAAGAAAGTAGGAAACCACTAGATCATCTCAGGTAAAGCTTAA 1327
Db 69542 AATTCAGACTTAAATTTGAAGAAAGTAGGAAACCACTAGACCGTTTCAGGTATGACCTTAA 69601
Qy 1328 ATCCAAATCCCTTATGATTATACAGTGGAAAGTAGAAATAGATTTAAGGCGCTAGATCTGA 1387
Db 69602 ATCCAAATCCCTTATGATTATACAGTGGAAAGTAGAAATAGATTTAAGGCGCTAGATCTGA 69661
Qy 1388 TAGACAGAGTACCTTAATGAATATGACAGAGTTTATGATGTCAGAGACAGGGA 1447
Db 69662 TACATAGAGTGGCTGATGAACATGGAATGAGTTTCTGATGATGTCAGAGACAGGGA 69721
Qy 1448 TCGACACCATCCCAATGGAAGAAATGCAAAAAGCAAAATGGCTGTCTGGGAGGCT 1507
Db 69722 TCAAGACCATCCCAATGGAAGAAATGCAAAAAGCAAAATGGCTGTCTGGGAGGCT 69781
Qy 1508 TACAATAGCTGTGAAAAGAGAGAAAGTGAAGCAAGGAAAGAAAGAAAGATAAAGC 1567
Db 69782 TACAACATGTTGTGAAAAGAGAGAAAGTGAAGCAAGGAAAGAAAGATAAAGC 69841
Qy 1568 ATCTCAATGTCAGAGTTCCAAA----- 1588
Db 69842 ATCTCAATGTCAGAGTTCCAAAAGATAGCAAGAGAAATAAGAAAGCCTTCCTCAGCGACC 69901
Qy 1589 ----- 1588
Db 69902 AGTGCAAAAGAAACAGAGGAAACAAATAGAATGGGAAGACTAGAGATCTCTTCAAAAAA 69961
Qy 1589 ----- 1588
Db 69962 TTAGAGATACCAAGGGAACATTTTCATGCAAGATGGGCTCGATAGAGACAGAAATGGTA 70021
Qy 1589 ----- 1588
Db 70022 TGCACCTAAAAGAGCTGACAATATTAAAGAGAGGTGGCAAGAAATACACAGAAAGCCGT 70081
Qy 1589 ----- 1588
Db 70082 ACAAAAAAGATCTTTTACAAACCAAGATATACAGATGGTGTGATCACTGACCTAGAGCCAG 70141
Qy 1589 ----- 1588
Db 70142 ACATCCTGGAACGTGAAGTCAAGTGGGCTTTAGAAAGCATCACTATGAACAAAGCTAGTG 70201
Qy 1589 ----- 1588
Db 70202 GAGGTGATGGAATTCAGTTGAGCTATTTCAAATCCTGGAAGATGATCCTGTGAAAGTGC 70261

Qy 1589 ----- 1588
Db 70262 TGCACCTCAATATGCCAGCAAAATTTGGAAAACTCAGCAGGGGCCACAGGACTGGAAAAAGT 70321
Qy 1589 ----- 1588
Db 70322 CAGTTTTCATTTCAATCCCAAGAAAGCAATGCCAAAGAAATGCTCAAACTACCAATAA 70381
Qy 1589 ----- 1588
Db 70382 TTGCACTCATCTCACACACTAGGAAGTAATGCTCAAAATTTCTCAAGCCAGGTTTCAGC 70441
Qy 1589 ----- GAACTTCCAGTTGTTTCAAGCTGGTTTTAGAAAAAGTCAAGAGAACCAAGACCA 1641
Db 70442 AATATGTGAATACCAAGTGTTTAAGCTGGTTTTAGAAAAAGGCAAGGAAACCAAGATCA 70501
Qy 1642 AATTGCCAATCCTCTGTATCATGGAAGAAAGCAAGAGAGTTCAGAAAAACATCTATTT 1701
Db 70502 AATTGCCAATCCTCTGTATCATGGAAGAAAGCAAGAGAGTTCAGAAAAACATCTATTT 70561
Qy 1702 CTGCTTTATGACTATGCAAAAGCCTTTGACTGTGGGCTCACAATAAACTGTGAAAAAT 1761
Db 70562 CTGCTTTATGACTATGCAAAAGCCTTTGACTGTGGGCTCACAATAAACTGTGAAAAAT 70621
Qy 1762 TCTGAAGGAGTGGGAATACCAAGCACCTGACCTGCTTCTGAAAAATTTGTATCGAG 1821
Db 70622 TCTGAAGAGATGGGAATACCAAGCACCTGACCTGCTTCTGAAAAATTTGTATCGAG 70681
Qy 1822 TCAGGAAGCAACAGTTAGAACTGGACATGGAAACAACAGAGCTGGTTCCAAAGTAGAAAAAG 1881
Db 70682 TCAGGAAGCAACAGTTAGAACTGGACATGGAAACAACAGAGCTGGTTCCAAAGTAGAAAAAG 70741
Qy 1882 AGTATGTCAGGCTGTATATGTCACCGGCTTGTAACTTCTATGCGAGAG-ACATCAT 1940
Db 70742 AGTATGTCAGGCTGTATATGTCACCGGCTTGTAACTTCTATGCGAGAGTACATCAT 70801
Qy 1941 GAGAAACGCTGGGCTGGGAAGAACACAGCTGGAATCAAGATTCGCGGAGAAATAGCAA 2000
Db 70802 GAGAAACGCTGGGAGAGAAACACAGCTGGAATCAAGATTCGCGGAGAAATAGCAA 70861
Qy 2001 TAACTTCAGATATGCGAGATGATACCACTTATGCGAGAAAGTGAAGAGAACTAAAAAG 2060
Db 70862 TAACTTCAGATATGCGATATGACACCACTTATGCGAGAAAGTGAAGAGAACTAAAAAG 70921
Qy 2061 CCTCTTGTGAAGGTGAAGAGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAA 2120
Db 70922 CCTCTTGTGAAGGTGAAGAGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAA 70981
Qy 2121 AACGAAGATCATGGCATCTGGTCCCATCACTTCATGGGAAATAGATGGGGAACAGTGA 2180
Db 70982 AATGAAGATCATGGTATCTGGTCCCATCACTTCATGGGAAAGAGATGGGGAACAGTGA 71041
Qy 2181 AACAGTGTGAGCTTTATTTTGGGGGCTCCAAATCACTGCAGATGGTGCAGC 2240
Db 71042 AACAGTGTGAGCTGTATTTTTT- TTGCTCTCCAAATCACTGCAGATGGTGCAGC 71100
Qy 2241 CATGAAATTAAGAGACACTTACTCTTGGAAAGAAAGTTA-- ACCAACCTAGATAGCAT 2298
Db 71101 CATGAAATTAAGAGTGCCTTACTCTTGGAGGAAGTTATGACCACTAGATAGCAT 71160
Qy 2299 TTGAAAGCAGAGACATTAACCTTGGCAACAAAGCCCATCTCTAGTCAAGGCTATGGTTTT 2358
Db 71161 TTCAAAAGCAGAGACATTAACCTTGGCAACAAAGTCTCTTCTAGTCAAGGCTGTGGTTTT 71220
Qy 2359 CCAGTGGTCTATGATGATGAGAGTTCGACTGTGAAGAAAGCTGAGCACTGAAGAAAT 2418
Db 71221 CCAGTGGTCTATGATGATGAGAGTTCGACTGTGAAGAAAGCTGAGCGCTGAAGAAAT 71280
Qy 2419 GATGCTTTTGAAGTGTGGTGTGGAGAGACTCTTTGAGAGTCCCTTGGAGTGCACAGGAG 2478
Db 71281 GATGCTTTTGAAGTGTGGTGTGGAGAGATCTTTCGAGTCCCTTGGAGTGCACAGGAG 71340

2479	TCCAACCACTCCATTCTCAAGGAGATCAGCCCTGGATTCTTTTGGAAAGGAATGATGCTA	2538
71341	TCCAAACCACTCCATCCTAAAGAGAGATCAGTCTGCTGCTTTCTTTTGGAAAGGACTTTTCTGTG	71400
2539	RAGCTTGAACCTCCAGTACTTTTGGCCACCTGATCAGAAGAGCTGACTCACTGGAAAAAGACC	2598
71401	RAGCTTGAACCTCCATATACTTTTGGCCACCTCATGCGAAGAGTTGACTCATTTGGAAAAAGACT	71460
2599	CTGATGCTGGAGGGATTTCGGGGCAGGAGGAGAGAGGGGACGACAGAGGATGAGATGGCTG	2658
71461	CTGATGCTGGAGGGATTTCGGGGCAGGAGGAGAGGGGACGACAGAGGATGAGATGGCTG	71520
2659	GATGGCATCACTGACTCCGATGCACTGAGTCTGGGTGAACTCTCTGGAGTTGGTGATGGAC	2718
71521	GATGGCTTCCCGACACAAATGGACATGATTTCAGTGAACCTCCGGAGTTGGTGATG--	71577
2719	AGGAGGCTGTCTTCGGCGGATTCATGGGTCACAAGAGGTTGGACAGCTTCAGCAAC	2778
71578	-GTGAGGGCTGGGTTACTTGCACCTCATGGGTTCACAAAGAGTCGAAACGACTGAGTGAC	71636
2779	TGAACCTGAACCTGAACCTGTACTGAAACCTTAGTA	2811
71637	TGAACCTGAACCTGAACCTGAGGTGCCAACATCATA	71669

RESULT 45
AC140093/c
LOCUS
DEFINITION
AC140093
Bos taurus clone RP42-262C3, WORKING DRAFT SEQUENCE, 8 ordered
pieces
AC140093
ACCESSION
AC140093.2 GI:29150354
VERSION
KEYWORDS
HTG: HTGS PHASE2: HTGS DRAFT.
linear
HTG 22-MAR-2003

ORGANISM	BOS taurus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae; Bovidae; Bovinae; Bos.
AUTHORS	1 (bases 1 to 155001) Akter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Cariaga, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J., Hashiguchi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., LaricP., Lee-Lin, S.-Q., Legaspi, R., Maduro, Q.L., Maduro, V.B., Margules, E.H., Masello, C., Maskeri, B., Mcbowell, J., Paquiguian, C., Pearson, R., Portnoy, M.E., Prasad, A., Reddi-Dugue, N., Schandler, K., Schueler, M.G., Sison, C., Stantirrop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L., Wecherby, K.D., Wiggins, L., Young, A., and Green, B.D.
TITLE	NISC Comparative Sequencing Initiative
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 155001) Green, E.D.
AUTHORS	Direct Submission
TITLE	Submitted (21-FEB-2003) NIH Intramural Sequencing Center, 8717
JOURNAL	Govement Circle, Gaithersburg, MD 20877, USA
REFERENCE	3 (bases 1 to 155001)

TITLE
 JOURNAL
 Direct Submission
 Submitted (22-MAR-2003) NIH Intramural Sequencing Center, 8717
 Grovmont Circle, Gaithersburg, MD 20877, USA
 COMMENT
 On Mar 22, 2003 this sequence version replaced gi:28460769.
 ----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: <http://www.nisc.nih.gov>
 Contact: nisczoo@nigri.nih.gov
 ----- Project Information
 Center project name: dds
 Center clone name: 263C03

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated

order and orientation of each sequence contig has been re-established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring contigs, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics -----

Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 153316 bases at least Q40
Consensus quality: 153884 bases at least Q30
Consensus quality: 154212 bases at least Q20
Insert size: 138000; agarose-fp
Insert size: 154301; sum-of-contigs
Quality coverage: 9.57x in Q20 bases; agarose-fp
Quality coverage: 8.56x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

*	1	719:	contig of 719 bp in length
*	720	819:	gap of unknown length
*	820	14217:	contig of 13398 bp in length
*	14218	14317:	gap of unknown length
*	14318	16197:	contig of 1880 bp in length
*	16198	16297:	gap of unknown length
*	16298	18432:	contig of 2135 bp in length
*	18433	18532:	gap of unknown length
*	18533	32260:	contig of 13728 bp in length
*	32261	32360:	gap of unknown length
*	32361	76744:	contig of 44384 bp in length
*	76745	76844:	gap of unknown length
*	76845	108615:	contig of 32771 bp in length
*	108616	109715:	gap of unknown length
*	109716	155001:	contig of 45286 bp in length.

103/16 13301: Concrey
Location/Qualifiers

```
1. .155001
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone_xp42-262C3"
/clone_lib="RP42"
```

```

1. 75001
/notes="clone overlaps with GenBank Accession Number
AC139631 clone RP42-10822 (center project name ddr)"
1. 7719
/notes="assembly_fragment
clone_end:17
vector_side:left"
820. 14217
/notes="assembly_fragment"
14318. 16197
/notes="assembly_fragment"
16298. 18432
/notes="assembly_fragment"
18533. 32260
/notes="assembly_fragment"
32361. 76744
/notes="assembly_fragment"
76845. 109615
/notes="assembly_fragment"
109716. 155001

```

/note="assembly_fragment clone_end:396 vector_side:right"							
Query Match 52.1%; Score 1861.5; DB 1; Length 155001; Best Local Similarity 75.0%; Pred. No. 0; Matches 2331; Conservative 0; Mismatches 215; Indels 561; Gaps 6;							
Qy	358	AAACAGTATGAAAGCAAAATGATAGGATAGTAAAGAGGAACCTCCCGAGTCAAGTAGGT	417	Db	121376	TATGACCTAAACTAAATCACTTAAAGATTATACAGTGAAGTGAAGAAATAGATTAAAGGC	121317
Db	122336	AAAAGTATGAAAGGCAAAATGATAGGATAGTAAAGAGGAACCTCCCGAGTCAAGTAGGT	122277				
Qy	418	GCCCATATGCTACTGGAGATCAGTGGAGAAATTAATCTCCAGAAAGATGAGAGATGGAG	477	Qy	1378	CTAGATCTGATAGACAGATACCTTAATGACTATGACAGAGGTTTATGACATGTACAG	1437
Db	122276	GCCCAATGTGCTACTGGAGATCAGTGGAGATATTAATCTCCAGAAAGATGAAAGGATGGAG	122217	Db	121316	CTAGATCTGATAGATAGTGCCTGATGAACCTATGAATGAGGTTCTGTGATATTGTACAG	121257
Qy	478	CCAAAGCAAAAGAAATACCCAGCTGTGGATGTGACTGGTGATATAAGCAAGGTCCGATGC	537	Qy	1438	GAGACAGGGATCGAGAGCCATCCCATGGAAGAAAGAAATGCAAAAAAGCAAAATGGCTGTCT	1497
Db	122216	CCAAAGCAAAACAATACCCAGCTGTGGTGTGACTGGTGATAGAGCAAGGTCCGATGC	122157	Db	121256	GAGACAGGGATCAAGACCATCCCATGGAAGAAAGAAATGCAAAAAAGCAAAATGGCTGTCT	121197
Qy	538	TGTAAAGAGCAATATGTCATAGGAACCTGGAAATGTACAGTCCATGAATCAAGGCAAAATTG	597	Qy	1498	GCGGAGCCCTTACAAATAGCTGTGAAAAGAGAGAAAGTGAAGAGCAAAAGGAAAAAGGAA	1557
Db	122156	TGTAAAGAGCAATATGTCATAGGAACCTGGAAATGTACAGTCCATGAATCAAGGCAAAATTG	122097	Db	121196	GCGGAGCCCTTACAAATAGCTGTGAAAAGAGAGAAAGTGAAGAGCAAAAGGAAAAAGGAA	121137
Qy	598	GAAATGGTCAAAACAGAGATGCAAGAGTGAATGTCAACATTTAGGAATCAGCGAACTA	657	Qy	1558	AGATAAAAGCATCTGAATGAGAGTTCCCAAGAACTTCCAGTTG-----	1601
Db	122096	GAAATGGTCAAAACAGAGATGCAAGAGTGAATGTCAACATTTAGGAATCAGCGAACTG	122037	Db	121136	AGATAAAGCATCTGAATTCAGAGTTCCCAAGAAATAGCAAGAGATGAAGAAAGCTTTC	121077
Qy	658	AAATGGATGGATGGGTGAATTTAACTCAGATGACCAATATATCTACTCTGGGGCAG	717	Qy	1602	-----TTCAAGTGGT	1612
Db	122036	AAATGGATGGATGGGTGAATTTAACTCAGATGACCAATATATCTACTCTGGGGCAG	121977	Db	121076	TTCAAGCAATCAATGCAAGAAATAGAGGAAAAACAACAGAAATGGAAGACTAGAGATCTC	121017
Qy	718	GAAATCCCTCAGAAGAAATGGAGTACCCATCATGTCAACAAAGAGTCCGAATGCGAGTA	777	Qy	1613	TTTAGAAAAGTCAAGAGGAAACCCAG-----	1637
Db	121976	GAAATCCCTCAGAAGAAATGGAGTACCCATCATGTCAACAAAGAGTCCGAATGCGAGTA	121917	Db	121016	TTAAGAAAATTAAGATATCAAGAGGAAACATTTTCATGCAAGATGGGCTCAATAAAGGAC	120957
Qy	778	CTTGATCCAGTCTCAAAAACGACAGATGATCTCTGTGTTTCCAAAGGCAAAACCATTC	837	Qy	1638	-----	1637
Db	121916	CTTGATCAAGCTCAAAAATGACAGATGATCTCTCTGTTTCCAAAGGCAAAACCATTC	121857	Db	120956	AGAAATGGTATGGACCTAACAGAAAGCAGAGATATTAAAGAGAGTGGCAAGATATACA	120897
Qy	838	AATATCAGATTAATCCAGATCTATGCCCAACCAAGTAATGTCTGAAGAGCTGAAGTTGAA	897	Qy	1638	-----	1637
Db	121856	AATATCAGATTAATCCAGATCTATGCCCAACCAAGTAATGTCTGAAGAGCTGAAGTTGAA	121797	Db	120896	GAAGAACTATACAAAAGAGATCTTCAGCAAGATAATCAAGATGTGTGATCACTGAC	120837
Qy	898	CGGTCTATGAAGACCTCAAGACCTTTTGAACCTAACACCCCAAAAAGATGTCTTCTC	957	Qy	1638	-----	1637
Db	121796	CAGTTCTCTGAAGACCTACAGACCTTTTGAACCTAACACCCCAAAAAGATGTCTTCTC	121737	Db	120836	CTAGACGACAGACATCTCGAATGTGAAGTCAAGTGGGCTTAGAAGCATCACTACGAAC	120777
Qy	958	ATTATAGGGGACTGGAATGCAAAAGTAGGAAGCAAGAAACACCTGGAGTAACAGGCAAA	1017	Qy	1638	-----	1637
Db	121736	ATTATAGGGGACTGGAATGCAAAAGTAGGAAGCAAGAAACACCTGGAGTAACAGGCAAA	121677	Db	120776	AAAGCTAGGGAGGTGATGGAATTCAGTGGAACTATTTCAAATCCTGAAGATGATGCT	120717
Qy	1018	TTTGGCCCTTGAATACGGAATGAAGCGGCAAGAGCTAATAGAGTTTTCGCAAGAAAT	1077	Qy	1638	-----	1637
Db	121676	TTTGGCCCTTGAATGAAGCGGCAAGAGCTAATAGAGTTTTCGCAAGAAAT	121617	Db	120716	GTGAAAGTGTGCACCTCAATATGCAAGCAAAATTTGGAAGAACTCAGCAGTGGCCACAGGAC	120657
Qy	1078	GCACTGGTCTATGCAAAACCCCTCTTCCAAACACACAGAGAGAGACTCTACATGGACA	1137	Qy	1638	-----	1637
Db	121616	GCACTGGTCTATGCAAAACCCCTCTTCCAAACACACAGAGAGAGACTCTACATGGACA	121557	Db	120656	TGGAAGAGGTCAATTTTCATTCCTCAATCCCAAGAAATGCTCAAACTACCGCACAAATTGCAC	120597
Qy	1138	TCACAGATGTCAACACCGAAATCAGATTGATTTATTTCTTTCAGCCAAAGATGGAGA	1197	Qy	1638	-----	1637
Db	121556	TCACAGATGTCAACACCGAAATCAGATTGATTTATTTCTTTCAGCCAAAGATGGAGA	121497	Db	120596	TCATCTCACACGCTAGTAAAGTAATGTCTCAAAATTTCTCAAGTCACTTTCAGCAATACG	120537
Qy	1198	AGCTCTATACAGTCAAGCAAAACCAAGACCGAGGCTTACTGTGCTCAGATCATGAATC	1257	Qy	1638	-----ACCAATTTGCCAAACATCTCTGTATC	1663
Db	121496	AGCTCTATACAGTCAAGCAAAACCAAGACCGGAGCTGCTGTGCTCAGATCATGAATC	121437	Db	120536	TGAACCATGAACCTCCAGATGTTCAAGCTGTTTGTAGAAAAGCCACATCCGCTGGATC	120477
Qy	1258	CTTATTCGCAATTCAGATTTAAATTGAAGAAATAGGGAACCACTAGATCACTCAGG	1317	Qy	1664	ATGGAAGAAAGCAAGAGAGTTCCAGAAAACATCTATTTCTGCTTTTATTGACTATGCAAAA	1723
Db	121436	CTTATTCGCAATTCAGATTTAAATTGAAGAAATAGGGAACCACTAGATCACTCAGG	121377	Db	120476	ATCGAAAAGCAAGAGAGTTCCAGAAAACATCTATTTCTGCTTTTATTGACTATGCAAAA	120417
Qy	1318	TAAGACCTAAATCCATCCCTTATGATTATATACAGTGGAGTGAAGATAGATTAAAGGC	1377	Qy	1724	GCCTTTGACTGTGGGGTCACAATTAACCTGTGGAAAATTTCTAAAAGGGATGGAAATACCA	1783
				Db	120416	GCCTTTGACTGTGGATCACAATAAATCTGCGGAATTTCTGAAGAGATGGAAATACCA	120357
				Qy	1784	GACCACTCACTGACTCTTTGAAAAATTTGTATGAGGTGAGGAGCAACAGTTAGAACT	1843
				Db	120356	GACCACTCACTGACTCTTTGAAAAACCTGTACGAGGTCAGGAGCAACAGTTAGAACT	120297
				Qy	1844	GGACATGGAAACACAGACTGTTTCCAAAGTAGGAAAGGATGATGTCAAGGCTGTATATG	1903

Db 120296 GGACATGCAACACAGACTGGTTCCAAATAGGAAAGAGTACGTCAAGGCTGTATATTG 120237
Qy 1904 TCACCCGGCTTGTTTAACTTCTATGCAGAG-ACATCATGAGAAACGCTGGGCTGGAAGAA 1962
Db 120236 TCACCCCTGCTTATTTAATCTTATATGACAGATACATCATGAGAAACACTGGGCTGGAAGAA 120177
Qy 1963 GCACAAGCTGGAATCAAGATTGCCGGGAGAAATAGCAATACCTCAGATATGCAGATGAT 2022
Db 120176 GCACAAGCTGGAATCAAGATTGCCGGGAGAAATATCAATAACCTCAGATATGCAGATGAC 120117
Qy 2023 ACCACCTTATGCGCAGAAAGTGAAGAGAACTAAAGAGCCTCTTGATGAAGGTCAAGAG 2082
Db 120116 ACCACCTTATGCGCAGAAAGTGAAGAGAACTAAAGAGCCTCTTGATGAAGGTCAAGAG 120057
Qy 2083 GAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAAAAGAGATCATGGCATCTGGT 2142
Db 120056 GAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAAAAGAGATCATGGCATCTGGT 119997
Qy 2143 CCCATCACTTCATGGGAATAGATGGGGAACAGTGGAAACAGTGTCAAGCTTTATTTTT 2202
Db 119996 CCCATCACTTCATGGGAATAGATGGGGAACAGTGGAAACAGTGTCAAGCTTTATTTTT 119937
Qy 2203 TGGGGGGTCCAAAATCACTGCAGATGCTGACTGCAGCCATGAAATTAAGAGACACTTAC 2262
Db 119936 --GGGGGCTCCAAAATCACTGCAGATGCTGACTGCAGCCATGAAATTAAGAGAGCCTTAC 119879
Qy 2263 TCCTTGGAGAAAGTTA--ACCACCTAGATAGCATATTCGAAAGCAGACACATTAACCT 2320
Db 119878 TCCTTGGAGAAAGTTATGACCAACCCAGATAGCATATTCGAAAGCAGACACATTAACCT 119819
Qy 2321 TGCCAAACAAAGCCCATCTAGTCAAGGCTATGGTTTTCCAGTGGTCAATGATGATGTG 2380
Db 119818 TGCCAAACAAAGCTTCATCTAGTCAAGGCTATGGTTTTCCAGTGGTGTGATGGATGTG 119759
Qy 2381 AGAGTTGAGCTGTGAAGAAAGCTGAGCACTGAGAAATTCGATGCTTTGAACTGTGGTGT 2440
Db 119758 AGAGTTGGGCTGTGAAGAGGCTGAGCACCAAGAAATGGATGATTTTGAAGCTGGTGT 119699
Qy 2441 GGAGAAAGCTCTTGAGAGTCCCTTGGACTGCAAGAGATCCAAACAGTCCATTCCTGAAG 2500
Db 119698 G--GAAAACCTTTGAGAGTCCCTTGGACTGCAAGAGATCCAAACAGTCCATTCCTGAAG 119641
Qy 2501 AGATCAGCCCTGGGATTTCTTTGGAAGGAATGATGCTAAAGCTGAACCTCAGTACTTTG 2560
Db 119640 AGATCAGCCCTGGGATTTCTTTGGAAGGTTGATGCTGAAGCTGAACCTCAGTACTTTG 119581
Qy 2561 GCCACTGATCAGAGAGTGTACTCACTGGAAGACCTGATGCTGGAGGGATTTGGGG 2620
Db 119580 GCCACTCAGAGAGAGTTAACTCATTGGAAGAGCTCTGATGCTGGAGGGATTAGGG 119521
Qy 2621 GCAGGAGAGAGGGGACGACAGAGATGAGATGGCTGGATGGCATCACTGACTGATGG 2680
Db 119520 GCAGGAGAGAGGGGACGACAGAGATGAGATGGCTGGATGGCATCACTGACTGATGG 119461
Qy 2681 AGTGAATCTGGTGAACTCCCTGGAGTGGTATGAGACAGAGGAGCCCTCTCTGGGGGA 2740
Db 119460 ACATGAGTTTGAGTGAATTCGGGAGTTGGTATGAGCAGGAGGAGCCCTGGTGTGCTGGGA 119401
Qy 2741 TTCATGGGCTCAAGAGTTGGACACCACTGAGCAACTGAGCAACTGAACCTGAACCTGTA 2800
Db 119400 TTCATGGAGTCAACAGAGTCCGACACGACTGAGCGAGTGAACCTGAACCTGAACCTG 119341
Qy 2801 AAACCTTAGTGTATATTAATCTAGAAATAGTAAATTCATATGATATTCAAAATATTT 2860
Db 119340 TGTACTGTTGACAGGAGAGTAAAGCTGGTATAGGCATTAATGGAAGAGCTCTGGAGTTTTT 119281
Qy -2861 CATAATGTTGGTTAAGATAATAGATTTTCAAAATGATTTTATCTT 2907
Db 119280 TCAAAAATACAAATAGAACTACCATATATACAGGAATTTCACTT 119234

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 13:00:13 ; Search time 650 Seconds
(without alignments)
3.920 Million cell updates/sec

Title: US-10-664-775-2

Perfect score: 3572

Sequence: 1 gtcaggagggcgagctga.....gcaacaacagcagaagctt 3572

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 0.5

Searched: 41 seqs, 356661 residues

Total number of hits satisfying chosen parameters: 82

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : rng2.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
C 1	1842.4	51.6	54842	1	ABS55699
C 2	1051.2	29.4	38793	1	ABS55700
C 3	1028.7	28.8	17141	1	AAD26486
C 4	1012.1	28.3	78056	1	ABV93701
C 5	1009	28.2	10492	1	ABV58843
C 6	966.804	27.1	23653	1	ABK14793
C 7	918.101	25.7	11288	1	ABV72459
C 8	917.7	25.7	10492	1	AAZ58843
C 9	885.401	24.8	78056	1	ABV93701
C 10	672	18.8	3370	1	AAQ79534
C 11	667.9	18.7	2830	1	AAT96051
C 12	619.909	17.4	54842	1	ABS55699
C 13	589.4	16.5	6620	1	ABV72460
C 14	557.9	15.6	23653	1	ABK14793
C 15	494.6	13.8	13723	1	AAV7439
C 16	413.4	11.6	908	1	ABN74558
C 17	406.2	11.4	1333	1	AAH21005
C 18	388.9	10.9	1813	1	AAD26492
C 19	380.6	10.7	967	1	ADE81111
C 20	372.6	10.4	419	1	ABX48667
C 21	364.7	10.2	451	1	ABX35679
C 22	363.2	10.2	436	1	ABX40926
C 23	360.3	10.1	432	1	ABX39925
C 24	357.1	10.0	433	1	ABX39519
C 25	353.3	9.9	433	1	ABX40182
C 26	348.7	9.8	1848	1	AAH21025
C 27	347.4	9.7	436	1	ABX41993
C 28	344.199	9.6	31412	1	ABQ75995
C 29	344.199	9.6	31412	1	ABQ93702
C 30	342.4	9.6	11034	1	AAD26491
C 31	335.8	9.4	428	1	ABX42989
C 32	333	9.3	406	1	ABX36143
C 33	326.3	9.1	1529	1	AAH21026

C 34	324	9.1	423	1	ABX43146	Bovine EST associa
C 35	321.6	9.0	410	1	ABX48669	Bovine EST associa
C 36	320.8	9.0	423	1	ABX40334	Bovine EST associa
C 37	320.4	9.0	409	1	ABX49143	Bovine EST associa
C 38	316.4	8.9	529	1	AAZ56452	Human immune/haema
C 39	313.4	8.8	425	1	ABX44380	Bovine EST associa
C 40	312.4	8.7	1435	1	AAH21034	Bovine-derived DNA
C 41	307.8	8.6	422	1	ABX38226	Bovine EST associa
C 42	306.2	8.6	3649	1	AAH50416	Calf acetyl cholin
C 43	299.8	8.4	336	1	ABX46049	Bovine EST associa
C 44	299.4	8.4	447	1	ABX46780	Bovine EST associa
C 45	298.6	8.4	815	1	AAD26490	Sheep alpha (1, 3)

ALIGNMENTS

RESULT 1

ABS55699/C

ID ABS55699 standard; DNA; 54842 BP.

XX

AC ABS55699;

XX

DT 08-JAN-2003 (first entry)

XX

DE Bovine Claudin-16 deficiency associated polynucleotide sequence #2.

XX

KW Bovine; Claudin-16 deficiency; type 2 mutation; ds.

XX

OS Bos sp.

PN JP2002238570-A.

XX

PD 27-AUG-2002.

XX

PF 14-FEB-2001; 2001JP-00037623.

XX

PR 14-FEB-2001; 2001JP-00037623.

XX

PA (CHIK-) CHIKUSAN GIJUTSU KYOKAI SH.

PA (KACH-) KACHIKU KAIRYO JIGYODAN SH.

XX (DOKU-) DOKURITSU GYOSEI HOJIN KACHIKU KAIRYO CE.

DR WPI; 2002-744759/81.

XX

PT Gene diagnosis of type 2 mutation for bovine claudin-16 deficiency, comprises checking for mutation in 1-161 position in claudin-16 sequence.

XX

PS Example 1; Page 9-28; 43pp; Japanese.

XX

CC The invention describes gene diagnosis of type 2 mutation for bovine claudin-16 deficiency. The method involves (a) getting a bovine nucleic acid sample; (b) subjecting the sample to a gene amplification reaction to give a nucleic acid fragment; (c) checking the presence of mutation in the nucleic acid fragment of step (b). The method is used for gene diagnosis of type 2 mutation for bovine claudin-16 deficiency. This sequence represents a polynucleotide from the bovine claudin-16 deficiency associated gene in which the type 2 mutation occurs

XX

SQ Sequence 54842 BP; 15478 A; 10714 C; 10200 G; 18336 T; 0 U; 114 Other;

Query Match 51.6%; Score 1842.4; DB 1; Length 54842;

Best Local Similarity 75.6%; Pred. No. 4.8e-05;

Matches 2295; Conservative 0; Mismatches 131; Indels 608; Gaps 7;

QY 358 AACCATATGCTACTGGAGATCAGTGGAGAAATAAATCTCCAGGTCAGTAGGT 417

DB 16137 AACATATGAAAGGCAAAATGATAGGATCTGAAAGAGGAAACTCCAGGTCAGTAGGT 16078

QY 418 GCCCATATGCTACTGGAGATCAGTGGAGAAATAAATCTCCAGGTCAGTAGGT 477

DB 16077 GCCCATATGCTACTGGAGATCAGTGGAGAAATAAATCTCCAGGTCAGTAGGT 16018

QY 478 CCAAGCAAAAAGAAATACCCAGCTGTGGATGTGACTGGTGATATAAGCAAGGTCGGATGC 537
Db 16017 CCAAGCAAAAAGAAATACACAGCTGGGAGTGTGACTGGTGATAGAAGCAAGGTCGGATGC 15958
QY 538 TGTAAAGCAATATTTGCATAGGAACCTGGATGTCAAGTCCATGATCAAGCAAAATTC 597
Db 15957 TGTAAAGCAATATTTGCATAGGAACCTGGATGTCAAGTCCATGATCAAGCAAAATTC 15898
QY 598 GAAGTGGTCAAAAGAGATGCCAAGAGTGAATCTCAACATTTCTAGGAATCAGCGAACTA 657
Db 15897 GAAGTGGTCAAAAGAGATGCCAAGAGTGAATCTCAACATTTCTAGGAATCAGCGAACTG 15838
QY 658 AAATGGACTGGAATGGGTGAATTTAACTCAGATGACCAATTTATCTACTACTGCGGCAG 717
Db 15837 AAATGGACTGGAATGGGTGAATTTAACTCAGATGACCAATTTATCTACTACTGCGGCAG 15778
QY 718 GAATCCCTCAGAAGAAATGGAGTAGCCATCATGTCAACAAAGAGTCCGAAATGCACTA 777
Db 15777 GAATCCCTCAGAAGAAATGGAGTAGCCATCATGTCAACAAAGAGTCCGAAATGCACTA 15718
QY 778 CTTGGATGCACTCTCAAAAACGACAGATGATCTCTGTTTGTTCAGAGCAAAACCAATTC 837
Db 15717 CTTGGATGCAATCTCAAAAATGACATATGATCTCTGTTTGTTCAGAGCAAAACCAATTC 15658
QY 838 AATATCAGATTAATCCAGTCTATGCCCCACCAAGTAACTGCTGAAGAGCTGAAGTTGAA 897
Db 15657 AATATCAGATTAATCCAGTCTATGCCCCACCAAGTAACTGCTGAAGAGCTGAAGTTGAA 15598
QY 898 CGGTCTATGAAGACCTACAAGACCTTTTAGAACTAAACACCCAAAAGAGATGCTCTCTC 957
Db 15597 CAGTTCATGAAGACCTACAAGACCTTTTAGAACTAAACACCCAAAAGAGATGCTCTCTC 15538
QY 958 ATTATAGGGGCTGAATGCAAAAGTAGGAAGCAAGAAACACCTGAGTAACAGGCAAA 1017
Db 15537 ATTATAGGGGCTGAATGCAAAAGTAGGAAGCAAGAAACACCTGAGTAACAGGCAAA 15478
QY 1018 TTTGGCCCTTGAATACGGAATGAAAGCAGGCAAGACTAATAGAGTTTGTCCAGAAAT 1077
Db 15477 TTTGGCCCTTGAATACGGAATGAAAGCAGGCAAGACTAATAGAGTTTGTCCAGAAAT 15418
QY 1078 GCATGTGTATAGCAAAACACCTCTTCCAAACACCAAGAGAACTCTACATGGAACA 1137
Db 15417 GCATGTGTATAGCAAAACACCTCTTCCAAACACCAAGAGAACTCTACATGGAACA 15358
QY 1138 TCACAGATGTTCACACCGAAATCAGATGATTAATTTCTTCAGCAGCAAGATGAGGA 1197
Db 15357 TCACAGATGTTCACACCGAAATCAGATGATTAATTTCTTCAGCAGCAAGATGAGGA 15298
QY 1198 AGCTCTATACAGTACGCAAAACCAAGACCAAGAGCTTACTGTGGCTCAGATCATGAATC 1257
Db 15297 AGCTCTATACAGTACGCAAAACCAAGACCAAGAGCTTACTGTGGCTCAGATCATGAATC 15238
QY 1258 CTTATTGCAAAATCAGACTTAAATTTGAAGATGAGGAAACCACTAGATCACTCAGG 1317
Db 15237 TTTATTGCAAAATCAGACTTAAATTTGAAGATGAGGAAACCACTTAAACCAATTCAGG 15178
QY 1318 TAAGACCTTAAATCCAAATCCCTTATGATTTATACAGTGAAGTGAAGAAATAGATTTAAGGGC 1377
Db 15177 TATGACCTTAAATCAATCCCTTATGATTTATACAGTGAAGTGAAGAAATAGATTTAAGGGC 15118
QY 1378 CTAGATCTGATAGACAGATGACTTAATGAATCTAGGACAGAGTTTCATGACATTTGTACAG 1437
Db 15117 CTAGATCTGATAGACAGATGACTTAATGAATCTAGGACAGAGTTTCATGACATTTGTACAG 15058
QY 1438 GAGACAGGATCGACACCATCCCATGAAAGAAATGCAAAAAGCAAAATGCTGTCT 1497
Db 15057 GAGACAGATCTGACACCATCCCATGAAAGAAATGCAAAAAGCAAAATGCTGTCT 14999
QY 1498 GGGAGGCTTACAAATAGCTGTGAAAAGAGAGAGTGAAGCAAGGAAAGGAA 1557
Db 14998 GGGAGGCTTACAAATAGCTGTGAAAAGAGAGAGTGAAGCAAGGAAAGGAA 14939
QY 1558 AGATAAAGCATCTGAATGCAGAGTTCCAAA----- 1588

Db 14938 AGATCAAGCATCTGAATGCGAGGTTCCAAAGAAATAGCAAGAGATAAGAAAGCCTTC 14879
QY 1589 ----- 1588
Db 14878 CTCAGTGATCAATGCAAAAGAAATAGAGGAWCWCAGAAATGGGAAAGACTAGANATCTC 14819
QY 1589 ----- 1588
Db 14818 TTCAAGAAAATTAGAGATACCAAGGGAATATTTTCATGCAAGATGGGCTCGATAAAGGAC 14759
QY 1589 ----- 1588
Db 14758 AGAAATGGTATGGATCTAACAGACAGCAAGATATTAAGAGAGGTGGCAAGAAATACACA 14699
QY 1589 ----- 1588
Db 14698 GAAGAACTGGACAAAAGATCTTCATGACCCAGATAATCACGATGTTGTGATCAGTGCAC 14639
QY 1589 ----- 1588
Db 14638 CCAGAGCCAGAAATCCTGGAATGTGAAGTCAAGTGGGCTTAGAAAGCATCACATATGAAC 14579
QY 1589 ----- 1588
Db 14578 AAAGCTAGTGAAGTGAATCCAGTTCAGTTCAGTTCATTCAGTCTCTGAAGGATGATGCT 14519
QY 1589 ----- 1588
Db 14518 GTGAAAGTCTGCATCAATATGCGAGCAAAATTTGGAAAACCTCAGCAGTGGCCACAGGAC 14459
QY 1589 ----- 1588
Db 14458 TGGGAAAAGGTCACTTTTCATTCCTCAATCCCAAGAAAGCAATGCTCAAGAAATGCTCAAA 14399
QY 1589 ----- 1588
Db 14398 CTACCACACAATGCACCTCATCTCACACACTAGTAAAGTAATGCTCAAAATTTCTCAAGCC 14339
QY 1589 -----GAACTCCAGTTGTTCAAGCTGGTTTGTAGAAAAGTC 1624
Db 14338 AGCTTCAGCAATAGCTGAACCGTGAATTTCTGTATGTTCAAGCTGGTTTGTAGAAAAGGC 14279
QY 1625 AGAGGAACCAAGAGCAAAATTCGCAACATCTCTGTATGATGGAAGAAAGCAAGAGATTC 1684
Db 14278 AGAGGAACCAAGAGCAAAATTCGCAACATCTCTGTATGATGGAAGAAAGCAAGAGATTC 14219
QY 1685 CAGAAAACATCTATTTCTGCTTTATGACTATGCAAAAGCCTTTGACTGTGGGGGTCAAC 1744
Db 14218 CAGAAAACATCTATTTCTGCTTTATGACTATGCAAAAGCCTTTGACTGTGGGTCAAC 14159
QY 1745 AATAAACTGTGAAAATTTCTGAAAGGGATGGAATACCAGACCCTGACCTGACTCTTG 1804
Db 14158 AATAAACTGTGAAAATTTCTGAAAGGGATGGAATACCAGACCCTGACTCTGCTCTTG 14099
QY 1805 AAAATTTGTATGCGGTGAGGAACCAAGTGTAGAACTGGACATGGAACAGCAAGCTGG 1864
Db 14098 AGAAATTTGTATGCGGTGAGGAACCAAGTGTAGAACTGGACATGGAACCAAGCAAGCTGG 14039
QY 1865 TTCCAAGTAGGAAAAGAGGATGTGCAAGGCTGTATTTGTCAACCGGCTTGTGTTAACTTC 1924
Db 14038 TTCCAAATAGGAAAAGGAATACATCAAGCTGTGTATTGTCAACCTGCTTACATTA 13979
QY 1925 TATGCGAG-ACATCATGAGAAACGCTGGGCTGGGAAGAAAGCAAGCTGGAATCAAGATT 1983
Db 13978 TATGCGAGTACATCATGAGAAAAGGCTGCACTGGGAAGAGCAAGCTAGAAATCAAGATT 13919
QY 1984 GCGGAGAGAAATAGCAATAACCTCAGATATGCAGATGATACCACCTTATGCGAGAAAGT 2043
Db 13918 GCTGGGAGAAATATCAATTAACCTCAGTATGAGATGACACCACTTATGCGAGAAAGT 13859
QY 2044 GAAGAGAACTAAAAAGCCTCTTGATGAAGGTGAAAGAGAGTGAAGAAAGTTGGCTTA 2103

QY	1589	-----	1588
Db	33382	TCACATGACAAAGCTAGTGGAGGTGATGGAATTCAGTTGAGCTATTTCAAATCCTGA	33323
QY	1589	-----	1588
Db	33322	AAGATGATGCTGTGAAGTGTGCTGCAATCAATATGCCAGCAAAATTTGGAAATTCAGCAGT	33263
QY	1589	-----	1588
Db	33262	GCCACAGACTGGAAAGGTGAGTTTTCATCCCAATCTCAAGAAAGGCAATGCCAAGA	33203
QY	1589	-----	1588
Db	33202	ATGCTCAAACTACCGCACAAATTGMACTCATCTCACACACTAGTAAAGTAATGCTCAAAAT	33143
QY	1589	-----	1588
Db	33142	TCTCCAGCCAGGCTTCAACWMAATATGGGAACGTGGAACCTCTCTGTATCATGGAATAA	33083
QY	1613	TTTAGAAAGTCAGAGGAACAGAGACCAAAATTCGAACATCCCTCTGTATCATGGAATAA	1672
Db	33082	TTTAGAAAGGCGAAGAAACAGAGACCAAAATTCGAACATCCCTCTGTATCATGGAATAA	33023
QY	1673	GCAAGAGATTCAGAAATAATCTATTTCTGCTTTATTTGACTATGCAAAAGCCCTTGAC	1732
Db	33022	GCAAGAGAG-TCCAAAGAAACATCTATTTCTGCTTTATTTGACTATGCAAAAGCCCTTGAC	32964
QY	1733	TGTGGGGTCACAATAAATCTGGAATAATCTGAAAGGATGGGAATACACAGACCACTG	1792
Db	32963	TGTGGGATCAAGAAACTGTGGGAATTTCTGAAGAGATGGGAATACACAGACCACTG	32904
QY	1793	ACCTGACTCTTGAAATAATTTGTATCAGGTGAGGAACCAAGTGTAGAACTGGACATGGA	1852
Db	32903	ACCTGCTCTTGAGAAACCTATCTCAGGTGAGGAACCAAGTGTAGAACTGGACATGGA	32844
QY	1853	ACAACAGACTGTTCCAGTAGGAAAGAGTATGTCAAGCTGTATTTCTACCCGGC	1912
Db	32843	ACAACAGACTGTTTCAACAGGAAGAGTATGTCAAGCTGTATTTCTACCCGGC	32784
QY	1913	TTGTTTAACTTCTATCAGAGACATCATCAGAAACGCTGGCTGGGAAGCAACAAGCTG	1972
Db	32783	TTATTTAACTTATATGACAGTATCATATGAAAGCTGGCTGGGAAGCAACAAGCTG	32724
QY	1973	GAATCAAGATTCGGGAGAAATAGCAATTAACCTCAGATATGCAGATGATACCACTTA	2032
Db	32723	GAATCAAGATTCGGGAGAAATATCAATAACCTCAGATATGCAGATGATACCACTTCTTA	32664
QY	2033	TGGCAGAAAGTCAAGAGAACTTAAAGGCTCTTCATGAAGTGAAGAGAGAGTGA	2092
Db	32663	TGGCAGAAAGTCAAGAGAACTTAAAGGCTCTTCATGAAGTGAAGAGAGAGTGA	32604
QY	2093	AAGTTGGCTTAAAGCTCAACATTCAGAAACGAGATCATGGCATCTGGTCCCATCTT	2152
Db	32603	AATTTGGCTT-AAAGCTCAACATTCAGAAACGAGATCATGGCATCTGGTCCCATCTT	32545
QY	2153	CATGGGAATAGTGGGAAACAGTGGAAACAGTGTCAAGCTTTATTTTGGGGGGCTC	2212
Db	32544	CATGGGAATAGTGGGAAACAGTGGAAACAGTGTCAAGCTTTATTTTGGGGGGCTC	32487
QY	2213	CAAAATCACTGCAGATGTGACTGCAGGCATGAAATTAAGACACTTACTCTTTGGAG	2272
Db	32486	CAAAATCACTGCAGATGTGACTGCAGGCATGAAATTAAGACACTTACTCTTTGGAG	32427
QY	2273	AAAAGTTA--ACCACTAGATAGCATATTTGAAGACAGAGACATTTACTCTGCCAACA	2330
Db	32426	AAAAGTTATGACCAACCTAGATAGCATATTTGAAGACAGAGACATTTACTCTGCCAACA	32367
QY	2331	GGCCCATCTAGTCAAGGCTATGGTTTTTCCAGTGTGATGTGATGTGAGTGGAG	2390
Db	32366	GGTCCGCTAGTCAAGGCTATGGTTTTTCCAGTGTGATGTGATGTGAGTGGAG	32307
QY	2391	TGTGAAGAAAGCTGAGCACTGAGAAATTTGATGCTTTTGAACCTGCTTTCGGAAGACT	2450
Db	32306	TGTGAAGAAAGCTGAGCACTGAGAAATTTGATGCTTTTGAACCTGCTTTCGGAAGACT	32247
QY	2451	CTTGAGAGTCCCTTCGACTGCAAGGAGATCCAAACCACTCCATTTCTGAAGGAGATCAGCCC	2510
Db	32246	CTTGAGAGTTCCTTCGACTGCAAGGAGATCCAAACCACTCCATTTCTGAAGGAGATCAGCCC	32187
QY	2511	TCGGATTTCTTTGGAAGGAATGATCTAAAGCTGAAACTCCAGTACTTTTGGCCACCTGAT	2570
Db	32186	TCGGATTTCTTTGGAAGGAATGATCTAAAGCTGAAACTCCAGTACTTTTGGCCACCTGAT	32127
QY	2571	CAGAGAGCTGACTCACTCGAAAGACCTGATGCTGGAGGATTCGGGGCAGGAGGAG	2630
Db	32126	CGAAGAGTGTGACCATTTGGAAGAA--CTGATGCTGGAGGATTCGGGGCAGGAGGAG	32069
QY	2631	AAAGGGACACAGAGATGAGATGCTGATGGATCACTGATCGATGGAGTGAATCT	2690
Db	32068	GAGGGATGACAGAGATGAGATGCTGATGGATCACTGATCGATGGAGTGAATCT	32009
QY	2691	GGGTGAACCTCTGGAGTTGGTCTGATGACAGAGGAGGCTGCTCTGGGGATTCATGGGT	2750
Db	32008	GAATGAACCTCTGGAGTTGGTCTGATGACAGAGGAGGCTGCTCTGGGGATTCATGGGT	31949
QY	2751	CACAAAGAGTTGGACACGACTGAGCAACTGAACTGAACTGAACTGAACTGAACTGAACT	2810
Db	31948	TACAAAGAGTTCGACACACTACTGACGAGTGAACCTGAACTGAACTGAACTGAACTGAACT	31889
QY	2811	A 2811	
Db	31888	A 31888	
RESULT 3			
AAD26486/c			
ID	AAD26486 standard; DNA; 17141 BP.		
XX	AAD26486;		
AC	26-MAR-2002 (first entry)		
DT	Bovine gamma globin gene.		
XX	Bovine; xenotransplantation; Gal-alpha (1,3)Gal; GAl determinant;		
KW	alpha(1,3) galactosyltransferase; alpha (1, 3) GT; drug screening;		
KW	surgery; gene therapy; gamma globin gene; ds.		
OS	Bos sp.		
XX	WO200188096-A2.		
XX	22-NOV-2001.		
XX	14-MAY-2001; 2001WO-US015765.		
XX	15-MAY-2000; 2000US-0204148P.		
XX	13-JUN-2000; 2000US-00593316.		
XX	(GERO-) GERON CORP.		
XX	Denning C, Clark J;		
XX	WPI; 2002-089848/12.		
XX	New ovine tissue devoid of antibody-detectable Gal-alpha(1,3)Gal		
XX	determinants, for xenotransplantation, and in the treatment of the human		
XX	body by surgery or therapy.		
XX	Claim 18; Page 70-74; 86pp; English.		
XX	The patent discloses immunologically compatible animal tissue, suitable		
XX	for xenotransplantation into human patients. The invention also relates		
XX	to an ovine tissue devoid of antibody-detectable Gal alpha (1,3)Gal (GAL)		

CC determinants which are made by alpha (1,3) galactosyltransferase (GT). The
CC ovine tissue is useful for treatment of human body by surgery or therapy
CC and in xenotransplantation, by transplanting the ovine tissue into a
CC mammal having circulating antibody against Gal alpha(1,3)GT determinants.
CC Polynucleotide constructs of the invention are useful for inactivating an
CC alpha(1,3) GT gene in an ovine cell. Alpha (1,3) GT sequences are useful
CC for drug screening and for the production of GAL containing synthetic
CC oligosaccharides. Sequences of the invention are also useful in gene
CC therapy. The present sequence is bovine gamma globin gene
XX
SQ Sequence 17141 BP; 5078 A; 3416 C; 3372 G; 5275 T; 0 U; 0 Other;

Query Match 28.8%; Score 1028.7; DB 1; Length 17141;
Best Local Similarity 91.9%; Pred. No. 0.079;
Matches 1117; Conservative 0; Mismatches 93; Indels 5; Gaps 3;

QY 1599 GAATTCAGTGTTCAGCTGGTTTGTAGAAAGTCCAGAGAACCCAGAGCAAAATTCGCC 1648
Db 16837 GAATTCAGTGTTCAGCTGGTTTGTAGAAAGTCCAGAGAACCCAGAGCAAAATTCGCC 16778

QY 1649 AACATCTCTGTATCATGGAAGAACAGAGAGTTCAGAGAAACATCTATTTCTGCTTT 1708
Db 16777 AACATCTCTGTATCATGGAAGAACAGAGAGTTCAGAGAAACATCTATTTCTGCTTT 16718

QY 1709 ATTGACTATGCAAAAGCCTTTGACTGTGGGGTCCACAAATTAACCTGTGAAAATTCGAAA 1768
Db 16717 ATTGACTATGCAAAAGCCTTTGACTGTGGGGTCCACAAATTAACCTGTGAAAATTCGAAA 16658

QY 1769 GGGATGGGAATACAGACACCTGACCTGACTCTTTGAAAATTTGTATGCAGTTCAGAA 1828
Db 16657 GAGATGGGAATACAGACACCTGACCTGACTCTTTGAAAATTTGTATGCAGTTCAGAA 16598

QY 1829 GCAACAGTTAGAACTGACATGGAACCAACAGACTGTTCCAAAGTGGAAAAGGATGATGT 1888
Db 16597 GCAACAGTTAGAACTGACATGGAACCAACAGACTGTTCCAAAGTGGAAAAGGATGATGT 16538

QY 1889 CAAGCTGTATGTTCAACCGGCTTTGTTAACTTCTATGCAGAG-ACATCATGAGAAAC 1947
Db 16537 CAAGCTGTATGTTCAACCGGCTTTGTTAACTTCTATGCAGAGTACATCATGAGAAAC 16478

QY 1948 GCTGGGTGGAAGAGACAGCTGGAATCAAGATTGCCCGGAGAAATAGCAATACCTC 2007
Db 16477 GCTGGGTGGAAGAGACAGCTGGAATCAAGATTGCCCGGAGAAATAGCAATACCTC 16418

QY 2008 AGATATGCAGATGATACCACTTTATGCGAGAAAGTGAAGAGAACTTAAAGCCCTCTTG 2067
Db 16417 AGATATGCAGATGATACCACTTTATGCGAGAAAGTGAAGAGAACTTAAAGCCCTCTTG 16358

QY 2068 ATGAAGTGAAGAGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAACGAAG 2127
Db 16357 ATGAAGTGAAGAGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAACGAAG 16298

QY 2128 ATCATGGCATCTGGTCCCATCCTTCATGGAATAGATGGGAAACAGTGGAAACAGTG 2187
Db 16297 ATCATGGCATCTGGTCCCATCCTTCATGGAATAGATGGGAAACAGTGGAAACAGTG 16238

QY 2188 TCAGACTTTATTTTGGGGGCTCCAAAATTCATCTGCAGATGTTGATGCAGCCATGAAA 2247
Db 16237 GCAGACTTTATTTT--CTGGGCTCCAAAATTCATCTGCAGATGTTGATGCAGCCATGAAA 16180

QY 2248 TTTAAAGACATTTACTCTTGGAGAAAGTTA--ACCACTAGATAGCATATTTGAAA 2305
Db 16179 TTTAAAGACATTTACTCTTGGAGAAAGTTA--ACCACTAGATAGCATATTTGAAA 16120

QY 2306 GCAGAGACATTACCTTCCCAACAAAGCCCATCTAGTCAAGGCTATGTTTTCAGTGG 2365
Db 16119 GCAGAGACATTACCTTCCCAACAAAGGTTCTGTAGTCAAGGCTATGTTTTCAGTGG 16060

QY 2366 TCATGTATGATGAGAGTGGAGTGGAGTGAAGAAAGCTGAGCACTGAAGAAATTCGCTT 2425
Db 16059 TCATGTATGATGAGAGTGGAGTGGAGTGAAGAAAGCTGAGCACTGAAGAAATTCGCTT 16000

QY 2426 TTGAACCTGTGGTGTGGGAAGAGTCTTTGAGAGTCCCTTGGAGTCCCAAGGAGATCCAAAC 2485

Db 15999 TTGAACCTGTGGTGTGGGAAGAGTCTTTGAGAGTCCCTTGGAGTCCCAAGGAGATCCAAAC 15940
QY 2486 AGTCCATTCTGAAGAGAGATCAGCCCTGGGATTTCTTTGGAAGAAATGATGCTAAAGCTGA 2545
Db 15939 AGTCCATTCTGAAGAGAGATCAGCCCTGGGATTTCTTTGGAAGAAATGATGCTAAAGCTGA 15880
QY 2546 AACTCCAGTACTTTGGCCACCTCATGCGAAGAGTTGACTCATTTGGAAAAAGACTCTGATGC 2605
Db 15879 AACTCCAGTACTTTGGCCACCTCATGCGAAGAGTTGACTCATTTGGAAAAAGACTCTGATGC 15820
QY 2606 TGGGAGGATTTGGGCGAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2665
Db 15819 TGGGAGGATTTGGGCGAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 15760
QY 2666 TCACTGACTCGATGGAACGTCAGTCTCTGGGTGAACCTCTCTGAGTTGGTGTGACAGGAGG 2725
Db 15759 TCACTGACTCGATGGAACGTCAGTCTCTGGGTGAACCTCTCTGAGTTGGTGTGACAGGAGG 15700
QY 2726 CTTGCTCTCGGGGATTCATGGGGTCAAAAAGTGTGACACGACTGAGCAACTGAACTG 2785
Db 15699 CTTGCGGTCTGCAATTCATGGGGTGGCAAGAGTGGGACACGACTGAGCGACTGATCTG 15640
QY 2786 AACTGAACTGTACTG 2800
Db 15639 ATCTGATCTGATCTG 15625

RESULT 4
ABV99701
ID ABV99701 standard; cDNA; 78056 BP.
XX
AC ABV99701;
XX
DT 12-FEB-2003 (first entry)
XX
DE Bovine BSE-resistant prion protein cDNA.
XX
KW Transmissible spongiform encephalopathy; neuroprotective; prion protein;
KW bovine spongiform encephalopathy; transgenic; BSE; bovine; cervid; PrP;
KW TSE; gene; ss.
XX
OS Bos taurus.
XX
PN WO200279416-A2.
XX
PD 10-OCT-2002.
XX
PF 28-MAR-2002; 2002WO-US009652.
XX
PR 30-MAR-2001; 2001US-0280549P.
XX
PA (TEXA) UNIV TEXAS A & M SYSTEM.
XX
PI Dunne PW, Piedrahita J;
XX
DR WPI; 2003-092895/08.
XX
DR P-PSDB; ABP57900.
XX
PT New transgenic bovine and cervid useful for producing animals which are
PT resistant to bovine spongiform encephalopathy and transmissible
PT spongiform encephalopathy disease, comprise a transgene encoding a mutant
PT PrP polypeptide.
XX
PS Disclosure; Page 60-82; 98pp; English.
XX
CC The invention relates to a novel transgenic bovine/cervid comprising a
CC transgene encoding a mutant prion protein (PrP) polypeptide, in which a
CC substitution has been made at position 171 of the sequence, which renders
CC the bovine/cervid resistant to bovine spongiform encephalopathy (BSE) and
CC transmissible spongiform encephalopathy (TSE) disease, respectively. The
CC transgene of the invention has neuroprotective activity. The method is
CC useful for producing a transgenic bovine or cervid resistant to BSE and

CC TSE diseases. The bovine prion gene is useful for producing transgenic
CC cattle exhibiting resistance to bovine spongiform encephalopathy. The
CC sequence represents the transgene encoding the mutant bovine PrP
CC polypeptide
XX Sequence 78056 BP: 22052 A; 17462 C; 16384 G; 22258 T; 0 U; 0 Other;

Sequence 78056 BP: 22052 A; 17462 C; 16284 G; 22258 T; 0 U; 0 Other; 0 X

Query Match	28.3%;	Score	1012.1;	DB 1;	Length	78056;
Best Local Similarity	92.8%;	Pred. No.	0.02;			
Matches 1113;	Conservative	0;	Mismatches	79;	Indels	7;
Gaps	5;					
QY	1589	GAACCTCCAGTGTGTCAAGCTGGTTTTAGAAAAAGTCAGAGAAACACGAGACCAAAATGGCC	1648			
DB	33273	GAACCTCCTGATGTTCAAGGTGGTTTTAGAAAAGGCGAGAAACAGAGATCAAAATGGCC	33333			
QY	1649	AACATCCTCTGTATCATGGAAAAAGCAAGAGAGTTCCAGAAAAACATCTATTTCTGCTTT	1708			
DB	33333	AACATCTGCTCGGATCATGGAAAAAGCAAAAGAGTTCCAGAAAAACATCTATTTCTACTTT	33399			
QY	1709	ATTGACTATCCAAAAGCCTTTGACTGTGGGGTCCACAATAAATCTGTGAAAAATTCGAAA	1768			
DB	33393	ATTGACTATGCCAAGGCCCTTTGACTGTGTGATCACAATTAATCTGTGAAAAATTCGAAA	33455			
QY	1769	GGGATGGGAATCCAGACCACCTGACTGACTTTGAAAAATTTGTATGAGTCAAGTCAAGAA	1828			
DB	33453	GAGATGGGAATCCAGACCACCTGATCTGCTCTTTGAAAAATTTATATGAGGTCAAGAA	33511			
QY	1829	GCAACAGTTAGAACTGACATGGAACAACAAGACTGGTTTCCAAGTAGGAAAAAGAGATGT	1888			
DB	33513	GCAACAGTTAGAACTGACATGGAACAACAAGACTGGTTTCCAATAGGAAAAAGAGATGT	33577			
QY	1889	CAAGGCTGTATATTGTCCACCGGCTTGTAACTTTATGACAGACACATCATGAGAAAC	1947			
DB	33573	CAAGGCTGTATATTGTCCACCGGCTTGTAACTTTATGACAGATACATCATGAGAAAC	33633			
QY	1948	GCTGGGCTGGAAGAAGCACAACTGGAAATCAAGATTCGCGGGAATAATGCAATACCTC	2007			
DB	33633	GCTGGACTGGAAGAAACACAAGCTGGAAATCAAGATTCGCGGGAATAATGCAATACCTC	33699			
QY	2008	AGATATCCAGATGATACACCCCTTATGCGAGAAAGTGAAGAGAACTAAAAAGCCTCTTG	2067			
DB	33693	AGATATTCAGATGATACACCCCTTATGCGAGAAAGTGAAGAGAACTAAAAAGCCTCTTG	33755			
QY	2068	ATGAAGGTGAAAGAGAGAGTAAAAAGTTTGGCTTTAAAGCTCAACATTCAGAAAAACGAAG	2127			
DB	33753	AGGAAAGTGAAGGTGGAGAGTAAACAAGTTTGGCTTTAAAGCTCAACATTCAGAAAAACGAAG	33811			
QY	2128	ATCATGGCATCTGCTCCCATCATTATGCGGAATAGATGGGNAACAGTGGAAACAGTG	2187			
DB	33813	ATCATGGCATCTGCTCCCATCATTATGCGGAATAGATGGGNAACAGTGGAAACAGTG	33877			
QY	2188	TCACACTTTATTTTTTGGGGGGCTCCAAAATCACTGCAGATGGTGAAGTGCAGGCATGAAA	2247			
DB	33873	TCACACTTTATTTTTTCTGGGCTCCAAAATCACTGCAGATGGTGAAGTGCAGGCATGAAA	33933			
QY	2248	TTAAAAAGACACTTACTCTCTTGGAAAGAAAGTTA--ACCAACTAGATAGCATATTGAAAA	2305			
DB	33931	TTAAAAAGACGCTTCTCTCTTGGATGGAAGAGTTATGACCAACCTAGATAGCATATTCAAAA	33999			
QY	2306	GCAGAGACATPACCTTCGCCAACAAGCCCATCTAGTCAAGGCTATCGTTTTTCCAGTGG	2365			
DB	33991	-CAGAGAGCTTACTTTTGGCAACAAAGTTCTGCTGATGCAAGGCTATGG--TTTTCTGTGG	3404			
QY	2366	TCATGTATGATGTGAGAGTTGGACTGTGAAGAAAGCTGAGCACTGAAGAATTGATGCTTT	2425			
DB	34049	TCATGTATGATGTGAGAGTTGGACTGTGAAGAAAGGCTGAGCGCTGAAGAATTGATGCTTT	3410			
QY	2426	TTGAACCTGTGGTGTGGGAAGAAGCTCTTGAGAGTCCCTTGGACTGGAAGGAGATCCAAAC	2485			
DB	34109	TTGAACCTGTGGTGTGGGAAGAAGCTCTTGAGAGTCCCTTGGACTGGAAGGAGATCCAAAC	3416			
QY	2486	AGTCCATTCTGAAGGAGATCAGCCCTGGGATTTCTTTTGAAGGAATGATGCTAAAGCTGA	2545			

Db	34169	AGTCCATCCTGAAGACATCAGCCCTCGGGATTTC	TTCTTTGGAAGGAAATGATGCTATAAAGCTGA	34229
Qy	2545	AACTCCAGTACTTTGGCCACCTTGATCAGAAGAGCTGACT	CACTCACTGGAAAAGACCCCTGATGTC	2605
Db	34229	AACTCCAGTACTTTGGCCACCTCATGTGAAGAGTTGACT	CATTCATTTGGAAGAAGACTCTGATGTC	34289
Qy	2606	TGGAGGGAATTGGGGCAGGACGAGAAAGGGACGACAGAGGAT	GAGATGAGATGGCTGGATGGCA	2665
Db	34289	TGGAGGGAATTGGGGCAGGAGGAGAAGGGGACGACAGAGGAT	GATGATGGCTGGATGGCA	34348
Qy	2666	TCACTGACTCGATGGACGTGAGTCTGGGTGAACTCCTGGAGTT	TGTTGATGGACAGGGAGG	2725
Db	34349	TCACTGACTTGATGGATGTGAGTCTGATGTAAGTAACTCGG	GAGTTGGTCACTGACAGGGAGG	34408
Qy	2726	CTGTCTCTCGGCGATTCAATGGGGTCAAAAGAGTTGGACAG	CACTGACGAACTGAACT	2784
Db	34409	CTGTGCGTCTGCTGATTTCTATGGGCTGCGAAACAGTCGG	ACGTGATCTGACAGACTGATCT	34467

RESULT 5
AAZ58843/C
ID AAZ58843 standard: DNA: 10492 BP.

AA
AC
AAZ58843;

25-APR-2000 (first entry)

XX
DE Nucleotide sequence of bovine myostatin promoter-enhancer region.

AA Myostatin; promoter; muscle cell disorder; genetic marker; bovine;
KW
KW myogenic regulatory factor; oncogene; muscle growth;
KW muscular dystrophy gene; ss.

KW myogenic regulatory factor; ss. muscular dystrophy gene: ss.

AA
OS

XX	Key	Location/Qualifiers
FH		
FT	enhancer	9094. .9099

```

FF /tag=
FF "E-box 4"
FF /note=

```

```
FT          enhancer
5230. 5241
/*tag= b
```

FT enhancer . 9860. . 9865

FT /note="E-bo

```
FT      main_program  
/*tag= d
```

FT	enhancer	10095.	.1010
FT	enhancer	10095.	.1010

```
ET
/note= "E-bo
```

```

ET
ET
/*tag= f

```

FT	TATA_signal	10240.	.1024
----	-------------	--------	-------

ET /note= "TATA

```
ET
/*tag= h
```

FT	CDS	10403.	.1049

```
FT ...
/note= "part
```

PN WO200001810-A1.

PD 13-JAN-2000.

PF 07-JUL-1999; 99WO-NZ000107
VV

PR 07-JUL-1998; 98NZ-00330902
XX

PA (NZPA-) NEW ZEALAND PASTORAL

XX Bass JJ, Jeanplong F, Kambadur R, Sharma M;
PI
XX WPI; 2000-182112/16.
XX
XX Novel promoter sequences of the myostatin gene used to diagnose muscle
PT cell disorders.
XX
XX Claim 1; Fig 1; 32pp; English.
XX
XX The invention relates to a bovine myostatin promoter sequence. The
CC myostatin promoter polynucleotide sequences can be used in diagnostic
CC tests for muscle cell disorders, and as a genetic marker for the
CC selection, in particular, of cattle and sheep breeds that express low
CC levels of myostatin, and so have increased muscle mass. The promoter can
CC also be used for the transcriptional control of a heterologous gene of
CC interest, e.g. myogenic regulatory factors, myostatin and myostatin
CC receptor, oncogenes, genes that regulate muscle differentiation and
CC growth, the muscular dystrophy gene, and any other genes expressed in
CC muscle. The present sequence represents the nucleotide sequence of bovine
CC myostatin promoter-enhancer region
XX
XX Sequence 10492 BP; 3262 A; 2160 C; 2002 G; 3068 T; 0 U; 0 Other;
SQ

Query Match	28.2%;	Score 1009;	DB 1;	Length 10492;
Best Local Similarity	91.9%;	Pred. No. 0.15;		
Matches 1127;	Conservative 0;	Mismatches 85;	Indels 14;	Gaps 6;
QY	1585	CAAAAGAACTTCCAGTGTCTTCAAGCTGTGTTTTAGAAAAGTCAGAGGAACCCAGAGACCAAAAT	1644	
DB	5545	CCATGAACCTTCTGATGTTCAAGCTGTGTTTTAGAAAAGGCAGAGAACCCAGAGATCAAAAT	5486	
QY	1645	TGCCAACAATCCTCTGTATTCATATGGA AAAAGCAAGAGATTCACGAAAAACATCTATTTCCTG	1704	
DB	5485	TGCCAACAATCTGCTGGATCATCGAAAAAGCAAGAGATTCAGAAAAACATCTATTTCCTG	5426	
QY	1705	CTTTATTGACTATGCCAAAGCCTTTGACTGTGGGGTGCAGATAAATCTGTGGAAAATTCCT	1764	
DB	5425	CTTTATTGACTATGCCAAAGCCTTTGACTGTGTGGATCAATAAATCTGTCAAAATTCCT	5366	
QY	1765	GAAAGGGATG-----GGAAATACCAGACCACTGCACCTGACTCTTTGAAAAATTTGTAT	1816	
DB	5365	GAAAGAGATGTTCTGTACCACCTGACTGCTGCACCTGACCTCTTTTGAAAAACCTGTAT	5306	
QY	1817	GCAGGTCAGAAAGCAACAGTTAGAACTGGACATGGACACAGACACTGGTTCACAGTAGGA	1876	
DB	5305	GCAGGTCAGAAAGCAACAGTTAGAACTGGACATGGAACTGGAACACAGACTGGTTCACAAATAGGA	5246	
QY	1877	AAAGGAGTATGTCAAAGCTGTATATTGTCAACCGGCTGTGTTAACTTCTATGCAGAG-AC	1935	
DB	5245	AAAGGAGATGTCAAAGACTGTATTGTCAACCTGCTTATTAACTTCTATGCAGAGTAC	5186	
QY	1936	ATCATGAGAAACCTGGGCTGGAGAGACACAGCTGGATCAAGATTGCGGGAGAAAT	1995	
DB	5185	ATCATGAGAAACACTGGGCTGGAAAGACACAGCTGGAATCAAGATTGCGGGAGAAAT	5126	
QY	1996	AGCAATAACTCAGATATGCAGATGATACCACCCCTTATGGCAGAAAGTGAAGAGGAACATA	2055	
DB	5125	ATTAAATAACCTCAGATATGCAGATGACACACCCCTTATGCAGAAAGTGAAGAGGAACATC	5066	
QY	2056	AAAAGCCTCTTGATGAGGTGAAAGAGGAGTGA AAAAGTGGCTTAAAGCTCAACATT	2115	
DB	5065	AAAAGCCTCTTGATGAAAGTGAAGAGGAGATTGAAAAAGTTGGCTTAAAGCTCAACATT	5006	
QY	2116	CAGAAAAACGAAGATCATGGCATCTGGTCCCATCACTTTCATGGGAAAT-AGATGGGGAAC	2174	
DB	5005	CAGAAAAACTAAGATCATGGCATCCGGTCCCATCACTGCATGGGAAATPAAGATGGGGAAC	4946	
QY	2175	AGTGGAAACAGTCTCAGACTTTATTTTTGGGGGGCTCCAAAATCACTCAGATGTTGTAC	2234	
DB	4945	AGTGGAAACAGTGTCA-AGACTTTATTTTCTGGGGTCCAAAATCACTCAAAATGTTGTAC	4887	
QY	2235	TGCAGCCATGAAATTTAAAGACACTTACTCTCTTGGAGAAAAGTTA--ACCAACCTTAGAT	2292	

Db	4886	TGCAGCCATGAAATTAAGAGATGCTTACTCTCTGGAAGGAAGTTATGACCAACCTAGAT	4827
Qy	2293	AGCATATTGAAAGCAGAGACATTACCTTGCCACAAGCCCATCTAGTCAGAGGCTATG	2352
Db	4826	AGCATATTGAAAGCAGAGACATTACTTTGCCACAAGGTCCATCTAGTCAAGGCTATG	4767
Qy	2353	GTTTTTCCAGTGTGTCATGTATGGATGTGAGAGTTGGACTCTGAAGAAAGCTGAGCACTGA	2412
Db	4766	GTTTTTCTGTGGTTCATGTATGGATGTGAGAGTTGGACTCTGAAGAAAGCTGAGCGCGA	4707
Qy	2413	AGATTTCATGCTTTTGAACCTGTGGTGTGTGAGAGAGACTCTTGAGAGTCCCTTGGACTGCA	2472
Db	4706	AGAATTGATGCTTTTGAACCTGTGGTGTGTGAGAGAGACTCTTGAGAGTCCCTTGGACTGCA	4647
Qy	2473	AGGAGATCCAAACAGTCCATTCTTGAAGGAGATCAGCCCTGGGATTTCTTTGGAAGGAATG	2532
Db	4646	AGGAGATCCAAACAGTCCATTCTTGAAGGAGATCAGCCCTGGGATTTCTTTGGAAGGAATG	4587
Qy	2533	ATGCTAAAGCTGAAACTCCAGTACTTTGGGCCACTGATCAGAGAGCTGACTCACTGGAA	2592
Db	4586	ATGCTAAAGCTGAAACTCCAGTACTTTGGGCCACTGATGTAAGAGTTGACTCATTTGGAA	4527
Qy	2593	AAGACCCCTGATGCTGGGAGGGATTGGGGGCAGAGAGAGAGAGGCGACACAGAGACTGAGA	2652
Db	4526	AAGACTCTGATGCTGGGAGGGATTGGGGGCAGAGAGAGAGGCGACACAGAGACTGAGA	4467
Qy	2653	TGGCTGATGGCATCACTGATCGATGAGTCGTGAGTCTGGGTGAACTCTCTGGAGTTGGTG	2712
Db	4466	TGGCTGATGGCATCACTGATCGATGAGTCGTGAGTCTGGGTGAACTCTCTGGAGTTGGTG	4407
Qy	2713	ATGGA-CAGGAGGCGCTGCTCTGGCGGATTTCATGGGTCACAAAGAGTTGGACACGACT	2771
Db	4406	ATGACACAGAGGCGCTGGCTTCTGAGATTTCATGGGTGGCAAGATTGGACATGACT	4347
Qy	2772	GAGCACTGAACCTGAACCTGATA	2797
Db	4346	GAAGGACTGAACCTGAACCTGAA	4321
RESULT 6			
ABK14793/c			
ID	ABK14793 standard; DNA; 23653 BP.		
XX			
AC	ABK14793;		
XX			
DT	08-MAY-2002 (first entry)		
XX			
DE	Platelet-derived growth factor expression cassette insert used in pBC734.		
XX			
KW	Platelet-derived growth factor; PDGF; transgenic; milk; Gene;		
KW	wound healing; diabetic foot ulcer; decubitus ulcer; ophthalmic disease;		
KW	venous stasis ulcer; periodontal regeneration; bone formation;		
KW	prosthetic vascular graft; pBC734; ds.		
XX			
OS	Homo sapiens.		
XX			
FN	WO200198520-A1.		
XX			
PD	27-DEC-2001.		
XX			
PF	19-JUN-2001; 2001WO-US041044.		
XX			
FR	19-JUN-2000; 2000US-0212406P.		
XX			
PA	(GENZ) GENZYME TRANSGENICS CORP.		
XX			
PI	Echelard Y, Meade H, Eichner W, Sommermeyer K;		
XX			
DR	WPI; 2002-083329/11.		
XX			
PT	Production of platelet derived growth factor (PDGF) comprises expression		
PT	in the milk of a non-human transgenic animal.		

FT TATA_signal /note= "E-box 2"
 FT 9879..9885
 FT /tag= d
 FT /note= "TATA-box 3"
 FT 10095..10100
 FT /tag= e
 FT /note= "E-box 1"
 FT 10197..10201
 FT /tag= f
 FT /note= "CAAT-box"
 FT 10240..10246
 FT /tag= g
 FT /note= "TATA-box 2"
 FT 10284..10270
 FT /tag= h
 FT /note= "TATA-box 1"
 FT 10403..10492
 FT /tag= i
 FT /note= "partial coding fragment"
 XX
 PN WO200001810-A1.
 XX
 XX 13-JAN-2000.
 PD
 XX
 PF 07-JUL-1999; 99WO-NZ000107.
 XX
 XX 07-JUL-1999; 98NZ-00330902.
 PR
 XX
 XX (NZPA-) NEW ZEALAND PASTORAL AGRIC RES INST LTD.
 PA
 XX Bass JJ, Jeanplong P, Kambadur R, Sharma M;
 PI WPI; 2000-182112/16.
 XX
 DR Novel promoter sequences of the myostatin gene used to diagnose muscle
 PT cell disorders.
 FT
 PS Claim 1; Fig 1; 32pp; English.
 XX
 CC The invention relates to a bovine myostatin promoter sequence. The
 CC myostatin promoter polynucleotide sequences can be used in diagnostic
 CC tests for muscle cell disorders, and as a genetic marker for the
 CC selection, in particular, of cattle and sheep breeds that express low
 CC levels of myostatin, and so have increased muscle mass. The promoter can
 CC also be used for the transcriptional control of a heterologous gene of
 CC interest, e.g. myogenic regulatory factors, myostatin and myostatin
 CC receptor, oncogenes, genes that regulate muscle differentiation and
 CC growth, the muscular dystrophy gene, and any other genes expressed in
 CC muscle. The present sequence represents the nucleotide sequence of bovine
 CC myostatin promoter-enhancer region
 XX
 SQ Sequence 10492 BP; 3262 A; 2160 C; 2002 G; 3068 T; 0 U; 0 Other;
 Query Match 25.7%; Score 917.7; DB 1; Length 10492;
 Best Local Similarity 94.0%; Pred No. 0.3;
 Matches 995; Conservative 0; Mismatches 58; Indels 5; Gaps 3;
 QY 1741 TCACAATAAATCTCGAAATTTCTGAAAGGGGATGGGAATACAGACACACCTGACCTGACT 1800
 DB 5607 TCACAATAAATCTCGAAATTTCTGAAAGAGATGGGCATACAGACACACCGGATCTGCCT 5666
 QY 1801 CTTGAAATTTGTATGAGGTTCAGAGTCAGACACACCTGTAGACTGGACATGGAACAACAGA 1860
 DB 5667 CTTGAAATTTGTATGAGGTTCAGAGTCAGACACACCAATTTAGTGTGACATGGACACAGA 5726
 QY 1861 CTGGTTTCCAAAGTAGGAAAGAGTAGTGTCAAGGCTGTATATTGTACCCGGCTTTGTTAA 1920
 DB 5727 CTGGTTTCCAAAGTAGGAAAGAGGTTCGTCAAGGCTGTATATTGTACCCCTGTTTATTAA 5786
 QY 1921 CTTCTATGCAGAG-ACATCATGAGAACGCTGGGCTGGAAGAACACAGCTTGAATCAA 1979
 DB 5787 CTTCTATGCAGAGTACATCATGAGAACGCTGGGCTGGAAGAACACAGCTTGAATCAA 5846

QY 1980 GATTGCCGGAGAAATAGCAATAAAGCTCAGATATCAGATATACCAACCTTATGGCAGA 2039
 DB 5847 GATTCCGGAGAAATATCAATTAAGCTCAGATATCAGATATACCAACCTTATGGCAGA 5906
 QY 2040 AAGTGAAGAGGAACTAAAAAGCCTTTGATGAAGTGAAGAGAGAGTGAAGAAAGTTGG 2099
 DB 5907 AAGTGCAGAGGAACTAAAAAGCCTTTGATGAAGTGAAGAGAGTGAAGAAAGTTGG 5966
 QY 2100 CTTAAGCTCAACATTCAGAAAACAGATCATGCAATCTGCTCCCATCCTCATTCATGGA 2159
 DB 5967 CTTAATCTCAACATTCAGAAAACAGAGATCAAGCATCCGTCCTCATCATTCATGGA 6026
 QY 2160 AATAGATGGGAAACACAGTGGAAACAGTGTCAAGCTTTATTTTGGGGGCTCCAAATC 2219
 DB 6027 AATAGATGGGAAACACAGTGGAAACAGTGTCAAGCTTTATTTTCTGGGCTCCAAATC 6084
 QY 2220 ACTGAGATGGTACTGACGCCATGAAATTAAGACACACTTACTCTCTGGAGAAAGTT 2279
 DB 6085 ACTACAATGGTACTGACGCCATGAAATTAAGATGCTTACTCTCTGGAGAAAGTT 6144
 QY 2280 A--ACCAACCTAGATAGCATATTGAAAGCAGAGACATTACCTTGCCCAACAAAGCCCAT 2337
 DB 6145 ATGACCAACCTAGATAGCATATTGAAAGCAGAGACATTACCTTGCCCAACAAAGTCCAT 6204
 QY 2338 CTAGTCAAGGCTATGGTTTTCAGTGTGATGATGATGATGATGATGATGATGATGATGATG 2397
 DB 6205 CTAGTCAAGGCTATGGTTTTCAGTGTGATGATGATGATGATGATGATGATGATGATGATG 6264
 QY 2398 AAGCTGACACTGAAGATTCATGCTTTTGAACCTGTTGAGTGTGAGTGTGAGTGTGAGTGTG 2457
 DB 6265 AAGCTGAGTGCAGAAATTCATGCTTTTGAACCTGTTGAGTGTGAGTGTGAGTGTGAGTGTG 6324
 QY 2458 GTCCCTTGGACTGCAAGGAGATCCAAACAGTCCATTCCTGAAGGAGATCAGCCCTGGGATT 2517
 DB 6325 GTCCCTTGGACTGCAAGGAGATCCAAACAGTCCATTCCTGAAGGAGATCAGCCCTGGGATT 6384
 QY 2518 TCTTTGGAAGGATGATGCTTAAGCTGAACTCCAGTACTTTTGGCCACCTGATCAGAGA 2577
 DB 6385 TCTTTGGAAGGATGATGCTTAAGCTGAACTCCAGTACTTTTGGCCACCTGATCAGAGA 6444
 QY 2578 GCTGACTCACTGAAAGACCTGATGCTGGAGGATTTGGGGGAGGAGGAGGAGGAGGAGG 2637
 DB 6445 GCTGACTCACTGAAAGACCTGATGCTGGAGGATTTGGGGGAGGAGGAGGAGGAGGAGG 6504
 QY 2638 CGACAGAGATGAGATGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2697
 DB 6505 CGACAGAGATGAGATGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 6564
 QY 2698 CTCTCGAGTTGGTGTGATGACAGGAGGAGGCTGTCTGCGGCGATTTCATGGGCTCAAAAG 2757
 DB 6565 CTCTCGAGTTGGTGTGATGACAGGAGGAGGCTGTCTGCGGCGATTTCATGGGCTCAAAAG 6624
 QY 2758 AGTTGACACGACTGAGCAACTGAACTG 2785
 DB 6625 AGTTGACACGACTGAGCAACTGATCTG 6652

RESULT 9
 ABV99701/c
 ID ABV99701 standard; cDNA; 78056 BP.

XX AC ABV99701;
 XX DT 12-FEB-2003 (first entry)
 XX DE Bovine BSE-resistant prion protein cDNA.
 XX KW Transmissible spongiform encephalopathy; neuroprotective; prion protein;
 KW bovine spongiform encephalopathy; transgenic; BSE; bovine; cervid; PrP;
 XX TSE; gene; ss.
 OS Bos taurus.

PN WO200279416-A2.
XX 10-OCT-2002.
XX 28-MAR-2002; 2002WO-US009652.
XX 30-MAR-2001; 2001US-0280549P.
XX (TEXA) UNIV TEXAS A & M SYSTEM.
XX Dunne PW, Piedrahita J;
XX WPI; 2003-092895/08.
XX P-PSDB; ABP57900.
XX
XX New transgenic bovine and cervid useful for producing animals which are
PT resistant to bovine spongiform encephalopathy and transmissible
PT prioniform encephalopathy disease, comprise a transgene encoding a mutant
PT PrP polypeptide.
XX
XX Disclosure; Page 60-82; 98pp; English.
XX
XX The invention relates to a novel transgenic bovine/cervid comprising a
CC transgene encoding a mutant prion protein (PrP) polypeptide, in which a
CC substitution has been made at position 171 of the sequence, which renders
CC the bovine/cervid resistant to bovine spongiform encephalopathy (BSE) and
CC transmissible spongiform encephalopathy (TSE) disease, respectively. The
CC transgene of the invention has neuroprotective activity. The method is
CC useful for producing a transgenic bovine or cervid resistant to BSE and
CC TSE diseases. The bovine prion gene is useful for producing transgenic
CC cattle exhibiting resistance to bovine spongiform encephalopathy. The
CC sequence represents the transgene encoding the mutant bovine PrP
CC polypeptide
XX
XX Sequence 78056 BP; 22052 A; 17462 C; 16284 G; 22258 T; 0 U; 0 Other;
Query Match 24.8%; Score 885.401; DB 1; Length 78056;
Best Local Similarity 86.6%; Pred. No. 0.053;
Matches 1063; Conservative 0; Mismatches 151; Indels 14; Gaps 8;
QY 1573 AATCAGAGTCCCAAGAACTCCAGTGTTCAGCTGGTTTATAGAAAGTCAGAGAAC 1632
DB 1660 AAGTACATGACCAATGAACTCCAGATGTTCAAGCTGGATTAGAAAGCACAGGAAT 1601
QY 1633 CAGACACAAATTTGCAACATCCTCTGTATCATGGAAGAAAGCAAGAGAGTTCCAGAAAA 1692
DB 1600 CAGAGATCAATTTGCCAATCCATCTTGATCATGGAAGAAAGTAAGAGCGTTCCAGAAAA 1541
QY 1693 CATCTATTTCTTTATGTACTATGCAAAAGCCTTTGACGTGGGGGTCAATTAACCT 1752
DB 1540 CATCCATTTTTCCTTTATGTACTGTGCCAAAGCCTTTGACTGTGTGGATCAACAAACT 1481
QY 1753 GTGGAAATTTCTGAAGGGATGGGAATACCAGACCCTGACCTGACTCTTTGAAAAATTT 1812
DB 1480 GTGGAAATTTCTGAAGATGGGAATACCAGAGCCTGACCTGCTCTGAGAAATCT 1421
QY 1813 GTATCAGGTAGGAAGCAACAGTTAGAACTGGACATGGAACACAGACTGGTTCCAAAGT 1872
DB 1420 GTATGCAAGTTAGGAAGCAACAGTTAGAACTAGACATGGAACACATAGCTGGTTCCAAAT 1361
QY 1873 AGGAAGAGGTATCTCAGGCTGTATATTGTACCCGCTGTTTAACTTCTATGCGAGA 1932
DB 1360 CAGGAAGAGGTATCTCAGGCTGTATATTGTACCCCTGCTTATTTAATTATATGCGAGA 1301
QY 1933 G-ACATCATGAGAAACGCTGGCTGGAGAGAGCAACAGCTGGAATCAAGATTGCCGGAG 1991
DB 1300 GTACATCAGGAAATGCTGGACTGGAATAAGCAACAGCTGGAATTAAGTTTGGCAGGAG 1241
QY 1992 AATATAGCAATACCTCAGATATGAGATGATACCACTTATGGCAGAGAAAGTGAAGAGA 2051
DB 1240 AATATATCAATACTCAGATATGAGATGATACCACTTATGG--TTAAGAGAGAGAGA 1183
QY 2052 ACTAAAAAGCCTCTTTGATGAAGTGAAGAGGAGAGTGAAGAAAGTTGGCTTAAAGCTCAA 2111

DB 1182 CCTAAAGAGCCTTTTGTATGAAAGTGAAGAGAGAGTGAAGAAAGTTGACTTAAACTCAA 1123
QY 2112 CATTCAGAAAAACGAAGATCATGGCATCTGGTCCCATCACTTCTATGGGAAATAGATGGGA 2171
DB 1122 CATTCAGAAAAACGAAGATCATGGCATCTGGTCCCATCACTTCTATGGGAAATAGATGGGA 1063
QY 2172 AACAGTGGAAACAGTGTCAAGCTTTATTTTGGGGGGCTCCAAATCACTCTCAGATGT 2231
DB 1062 AACAGTGGAAACAGTGTCAAGCTTTATTTTGGGGGGCTCCAAATCACTCTCAGATGT 1006
QY 2232 GACTGCAGGCATGAATTAATAAGACACTTACTCTTTGGGAAGAAAAAGTTA--ACCAACCTA 2289
DB 1005 GACTGCAGGCATGAATTAATAAGTCTTGTCTTGGGAAGAAAAAGTTATGACCAACCTA 947
QY 2290 GATAGCATATTGAAGACAGACATTAATCTTCCCAACAAAGCCCATCTAGTCAAGCT 2349
DB 946 GACAGCATATTAAAAAGCAGACATTAATCTTCCCAACAAAGCTCCATCTAGTCAAGCT 887
QY 2350 ATGTTTCTCCAGTGGTGTATGATGTGAGAGTGGACTGTGAAGAAAGCTGAGCAC 2409
DB 886 ATGTTTCTCCAGTGTATGATGTGAGAGTGGACTGTGAAGAAAGCTGAGTGC 827
QY 2410 TGAAGAAATGATCTTTTGAACCTGGTGTGGAGAAAGCTCTTGAAGCTCCCTTGGACT 2469
DB 826 TGAAGAAATGATCTTTTGAACCTGGTGTGGAGAAAGCTCTTGAAGCTCATTGGACT 767
QY 2470 GCAAGAGATCCAAACAGTCCATCTCTGAAGGAGATCAGCCCTGGGATTTCTTTGGAAGA 2529
DB 766 GCA---AGATCAACACGCTCCATCTTAAAGGAAATCAGTCCCTGAATATTATCGGAAGA 710
QY 2530 ATGATGCTTAAAGCTGAAACTCCAGTACTTTGGCCACTTCAATCAGAAAGCTGACTCACTG 2589
DB 709 CTAAATGCTGAAGCTGAAAGTCCAAATCTTTGGCCACTGATGCAAGAAAGCTGCTCATTTG 650
QY 2590 GAAAGACCTGTATGCTGGAGGGATTTGGGCGCAGGAGAGAGGGGACGACAGAGGATG 2649
DB 649 GAAAGAAACCTGTATGCTGAGATGATTGAAGTGGGAGGAGAGGGGACGACAGAGATG 590
QY 2650 AGATGGCTGGATGGCATCACTGATCGATGGAGCTGAGTCTGGGTGAACCTCTCTGGAGTTG 2709
DB 589 AGATGGTGGATGGCATCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 530
QY 2710 GTATGGAC--AGGAGGCTGTCTCTCGGGGATTTATGGGGTCAAAAGAGTTGGACAG 2768
DB 529 GTATGGACAAAGGAGGAGGCTGGCATGCTGAGTCCATGGGGTTGCAAGAG--TCGATACG 471
QY 2769 ACTGAGCAACTGAACCTGAACCTGAACCTGT 2796
DB 470 ACTGAGCACTGAACCTGAACCTGAACCTGT 443
RESULT 10
AAQ79534/C
ID AAQ79534 standard; DNA; 3370 BP.
XX AAQ79534;
AC AAQ79534;
XX
XX 25-MAR-2003 (revised)
DT 04-AUG-1995 (first entry)
XX
XX Bovine tracheal antimicrobial peptide gene.
DE Antimicrobial peptide; TAP; ss.
XX
XX Bos taurus.
XX
XX Key Location/Qualifiers
FT intron 1..1472
FT /*tag= e
FT CAAT_signal 1382..1389
FT /*tag= a
FT TATA_signal 1444..1449

```

FT      exon      /*tag= b
FT      FT      1503..1560
FT      /tag= c
FT      intron    1561..1073
FT      /tag= f
FT      exon      3074..3211
FT      /tag= d
FT      intron    3212..3370
FT      /tag= g
XX      WO9426106-Al.
XX      PD
XX      24-NOV-1994.
XX      PF      11-MAY-1994; 94WO-US005257.
XX      PR      11-MAY-1993; 93US-00060822.
XX      (CHIL-) CHILDRENS HOSPITAL PHILADELPHIA.
XX      PI      Zasloff MA, Bevins CL, Diamond G;
XX      WPI; 1995-006233/01.
XX      DR      P-PSDB; AAR66205.
XX      PT      New antimicrobial peptide and precursor from mammalian trachea - and
XX      PT      related DNA, vector and transformed cells, useful as pharmaceuticals and
XX      PT      disinfectants.
XX      PS      Claim 12; Page 40-42; 69pp; English.
XX      CC      The tracheal antimicrobial peptide (TAP) gene is expressed primarily in
XX      CC      the columnar epithelial cells of the mucosa lining the bovine airway. RNA
XX      CC      analysis indicates that TAP is expressed along the entire length of the
XX      CC      conducting airway. It includes a consensus sequence for an NFkB
XX      CC      recognition site 181 bases upstream from the transcriptional start site.
XX      CC      When assayed in vitro several different strains of microbes, including
XX      CC      some which are respiratory pathogens, a mammalian TAP (see AAR66204 FT)
XX      CC      of bovine origin showed similar inhibitory activity to that of synthetic
XX      CC      marainin 2-NH2. While TAP (see AAR66204 FT) was most active against E.
XX      CC      coli and K. pneumoniae, significant antimicrobial activity was also seen
XX      CC      when applied to C. albicans. (Updated on 25-MAR-2003 to correct PN
XX      CC      field.)
XX      SQ      Sequence 3370 BP; 826 A; 872 C; 811 G; 861 T; 0 U; 0 Other;

Query Match      18.8%; Score 672; DB 1; Length 3370;
Best Local Similarity 89.1%; Pred. No. 6.1;
Matches 778; Conservative 0; Mismatches 85; Indels 10; Gaps 5;

QY      1950 TGGGCTGGAAGAAGCACAAAGCTGAATCAAGATTGCCGGGAGAAATAGCAATACCTCAG 2009
DB      939 TTGGCTGGAAGACACAGCTGGATCAAGATTGCCGGGAGAAATGTCAATACCTCAA 880
QY      2010 ATATGCAGATGATACCACTTTATGGCAGAAAGTGAAGAGAACTAAAAGCCCTTCAT 2069
DB      879 ATATGCAGATGATACCACTTTATGGCAGAAAGTGAAGAGAACTAAAAGCCCTTCAT 820
QY      2070 GAAAGTGAAGAGGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAAACGAAGAT 2129
DB      819 GAAAGTGAAGAGTGAAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAAACGAAGAT 760
QY      2130 CATGCCATCTGGTCCCATCATCTCATGGGAAATAGATGGGAAACAGTGGAAACAGTGTC 2189
DB      759 AATGGCGTATGATCCATCACTCATGCGAAATAGATGGGAAACAGTGGAAACAGTGTC 700
QY      2190 AGACTTTATTTTGGGGGGCTCAAAATCACTGCAGATGGTGAATGCAGCCATGAATT 2249
DB      699 AGACTTTATTTT---GGGGCTCAAAATCACTGCAGATGGTGAATGCAGCCATGAATT 643
QY      2250 AAAAGACACTTACTCTTTGGGAGAAAGTTA--ACCAACCTAGATAGCATATTGAAAAGC 2307
DB      642 AAAACACGATTACTCTTTGGGAGAAAGTTATATGACCAACCTAGATAGCAT--TTGAAAAGC 584

```

RESULT 11

AAT96051

ID AAT96051 standard; DNA; 2830 BP.

XX AC AAT96051;

XX XX 01-MAY-1998 (first entry)

XX DE DNA for bovine brain protein p97.

XX XX Bovine; brain protein; p97; genetic engineering; analysis; ds.

XX OS Bos taurus.

XX PH Key Location/Qualifiers

XX FT CDS 123..1901

XX FT /*tag= a

XX FT /product= "p97"

XX PN JP10001497-A.

XX XX 06-JAN-1998.

XX PF 13-JUN-1996; 96JP-00152684.

XX PR 13-JUN-1996; 96JP-00152684.

XX XX (MITU) MITSUBISHI CHEM CORP.

XX XX WPI; 1998-114806/11.

XX DR P-PSDB; AAW38337.

XX PT New protein and a DNA encoding it - useful in genetic engineering for

XX PT various analyses.

```
PS Claim 3; Page 6-8; 12pp; Japanese.
XX The present sequence encodes the bovine brain protein p97, which is
CC useful in genetic engineering techniques for carrying out various
CC analyses
XX
SQ Sequence 2830 BP; 938 A; 522 C; 686 G; 684 T; 0 U; 0 Other;

Query Match      18.7%; Score 667.9; DB 1; Length 2830;
Best Local Similarity 81.7%; Pred. No. 7.4;
Matches 782; Conservative 0; Mismatches 166; Indels 9; Gaps 1;

Qy 369 AAGCAAAATCATAGGATCTGAAGAGGAACTCCCGAGTCTAGTAGTCCCGCATATGC 428
Db |||||
Qy 659 AAGGCTCTCTGGCAGGACATCGAAAGATGAGCCCGCAGGTCCGAAAGGGTTCAACATGC 718
Db |||||
Qy 429 TACTGGAGATCAGTGGAGA-----AATACTCCAGAAAGAAATGAAGAGATGGAGCC 479
Db |||||
Qy 719 TACTGGAGAGAGCGGAGGCGCGATCTAATACCTCCAGTAAGAAATGAAGCGGCTGGCA 778
Db |||||
Qy 480 AAAGCAAAAGAAATACCAGTCTGGATGTGACTGTGTATATAGCAAGTCCGATGCTG 539
Db |||||
Qy 779 AAAGTGAAGAGGACAGTCAGCTGTGATGTCTGTGGTGAAGAAATGAATCCGATGCTG 838
Db |||||
Qy 540 TAAAGAGCAATATTGCCATAGGAACCTGGAATGTCCAGTCCATGAATCAAGGCAAAATTGGA 599
Db |||||
Qy 839 TAAAGAGCAATCTGCATAGGAACCTGGAATGTTAGATCTATGATCTCTGTAATTTGGA 898
Db |||||
Qy 600 AGTGGTCAAAAGAGATGGCAAGTGAATGTCAACATCTTAGGAATCAGCGAACTAAA 659
Db |||||
Qy 899 TGTGGTGAAGCAGGAGATGGAAGAAATAAATCATCGATCTTAGGAATCAGTGAACATAA 958
Db |||||
Qy 660 ATGCACTGGAATGGTGAATTTAATCTAGATGACCATTAATATCTACTACTCGGGCAGGA 719
Db |||||
Qy 959 ATGACAGGAATGGCGAATTTGAATTCAGATGACATTAATATCTACTACTGTGGGCAACA 1018
Db |||||
Qy 720 ATCCCTCAGAAAGATGAGTAGCCATCATGGTCAACAAAGAGTCCGAAATGCAGTACT 779
Db |||||
Qy 1019 ATCCCTTAGAAGAAATGAGTCTCATAGTTAAACAAAGAGTCCGAAATGCATAAT 1078
Db |||||
Qy 780 TGGATGAGTCTCAAAAGCAGAGATGATCTGTGTTGTTTCCAGGCAACCATTCAA 839
Db |||||
Qy 1079 TGGTGTCAATCTGAAACAGCAGATGATTTGAGTTTCCAGGCAACCATTCAA 1138
Db |||||
Qy 840 TATCACAGTAATCAAGTCTATGCCCAACAGTAATGCTGAAGAAGCTGAAGTTGAACG 899
Db |||||
Qy 1139 CCTCACAGTAATCCAAGTCTATGCCCACTCTTATGCTGAAGAAGTGAAGTTACCG 1198
Db |||||
Qy 900 GTCTATGAGACCTACAGACCTTTTGAACCTAACACCCCAAAAGATGCTCTCTCAT 959
Db |||||
Qy 1199 GTTCTATGAGACCTACAAACCTTCTGGAATAAACACCGGAAATAGATGCTCTTTTCTAT 1258
Db |||||
Qy 960 TATAGGGGACTGGAATGCAAAAGTAGGAAGCAAGAAACACCTGGAGTAAACAGGCAAAAT 1019
Db |||||
Qy 1259 CATAGGGGATTTGGAATGCAAAAGTGGGAGTCAAGAGATACCTGGAATACAGGACAGTT 1318
Db |||||
Qy 1020 TGGCCTTGGAAATACGGAATGAAGCAGGCAAGACTAATAGAGTTTTCGCAAGAAATGC 1079
Db |||||
Qy 1319 TGGCCTTGGAAATGCAAAATGAAGCAGGCGGAAGCTAATCGAGTTTGTCCACCAACAG 1378
Db |||||
Qy 1080 ACTGGTCTATAGCAACACCTCTTCCACACACAGAGAGACTCTACACATCGGATC 1139
Db |||||
Qy 1379 GCTGGTCTATAACAAACACCTTTTCCAAACACCTAGTAGAGCTCTACACATGGACATC 1438
Db |||||
Qy 1140 ACCAGATGGTCAACACCGGAAATCAGATTGATATATTTTTCGACGCAAAAGATGGAGAAG 1199
Db |||||
Qy 1439 ACCAGATGGTCTGATACCGAGATCAGATTGATTATATTATTGTCGCCAAAGATGGAGAAG 1498
Db |||||
Qy 1200 CTCTATACAGTCAAGCAAAACAAAGACAGGAGCTTACTGTGGCTCAGATCATGAACTCCT 1259
Db |||||
Qy 1499 CTCTGTACAGTCAAGCAAAACAAAGACCTGGAGCTGACTGTGGCTCAGATCATAGCTCCT 1558
Db |||||
Qy 1260 TATTGCCAAATTCAGACTTAAATTTGAAGAAAGTAGGGAACCACTAGACTCACTCAG 1316
Db |||||

Db 1559 TATTGCAAGTTTCAGGCTTAAGTTGAAGATAATACAAAGACACTCGGCCATTGAG 1615
|||
RESULT 12
ABS55699
ID ABS55699 standard; DNA; 54842 BP.
XX
AC ABS55699;
XX
DT 08-JAN-2003 (first entry)
XX
DE Bovine Claudin-16 deficiency associated polynucleotide sequence #2.
XX
KW Bovine; Claudin-16 deficiency; type 2 mutation; ds.
XX
OS Bos sp.
XX
PN JP2002238570-A.
XX
PD 27-AUG-2002.
XX
PF 14-FEB-2001; 2001JP-00037623.
XX
PR 14-FEB-2001; 2001JP-00037623.
XX
PX (CHIK-) CHIKUSAN GIJUTSU KYOKAI SH.
PA (KACH-) KACHIKU KAIRYO JIGYODAN SH.
PA (DOKU-) DOKURITSU GYOSEI HOJIN KACHIKU KAIRYO CE.
XX
DR WPI; 2002-744759/81.
XX
PT Gene diagnosis of type 2 mutation for bovine claudin-16 deficiency,
comprises checking for mutation in 1-161 position in claudin-16 sequence.
XX
PS Example 1; Page 9-28; 43pp; Japanese.
XX
CC The invention describes gene diagnosis of type 2 mutation for bovine
claudin-16 deficiency. The method involves (a) getting a bovine nucleic
acid sample; (b) subjecting the sample to a gene amplification reaction
to give a nucleic acid fragment; (c) checking the presence of mutation in
the nucleic acid fragment of step (b). The method is used for gene
diagnosis of type 2 mutation for bovine claudin-16 deficiency. This
sequence represents a polynucleotide from the bovine claudin-16
deficiency associated gene in which the type 2 mutation occurs
XX
SQ Sequence 54842 BP; 15478 A; 10714 C; 10200 G; 18336 T; 0 U; 114 Other;

Query Match      17.4%; Score 619.909; DB 1; Length 54842;
Best Local Similarity 63.9%; Pred. No. 0.56;
Matches 1166; Conservative 0; Mismatches 561; Indels 99; Gaps 16;

Qy 1656 TCTGTATCATGGAAGCAAGAGAGTTCAGAAAAACATCTATTCTGCTTTATTGACT 1715
Db |||||
Qy 9575 TCTGGATCATAGGAAGCAAGGGAATTCCA-AAGCACATCTATTCTACTTCATT-ATT 9632
Db |||||
Qy 1716 ATGCAAAAGCCTTTGACTGTGGGGGTCACAATAAATCTGGAATAATCTGAAAGGATGG 1775
Db |||||
Qy 9633 ACATAAAGCCTTTGCAATGTGGATCAGAAAAAATGAGAAATT-----AAAATTGG 9685
Db |||||
Qy 1776 GAATACCAAGACCTGACTGCTGCTCTTGAATAATTTGTATGCAGTCAAGGAAGCAACAG 1835
Db |||||
Qy 9686 AATATCCAGACACCTTACCTGCTCTCTGAGAAACCTGTTTTCAGGACAAAGATCAACAG 9745
Db |||||
Qy 1836 TTAGAACTGGACATGGAACAACAGACTGTTTCCAAAGTAGGAAAAAGAGTATGTCAAGGT 1895
Db |||||
Qy 9746 TTAGAACTGGACATGGAACAACAGACTGTTTCCAAATAGGAGAGAAATACGTCAGGCT 9805
Db |||||
Qy 1896 GTATATTGTCAACCGGCTTGTATTAACTTCTATGACAGACATCATCAGAAACGCTGGCT 1955
Db |||||
Qy 9806 GTATATTGTCACTTCTTATTAACTTATATTTCAGAACTACTTAT----- 9850
Db |||||
Qy 1956 GGAAGAAGCAACAGCTGGAATCAAGATTGCCGGGAGAAATAGCAATAAATCACTCAGATATGC 2015
Db |||||
```

```
Db 9851 -----ATGCGTATCTTATTTAAACGATATGC 9876
QY 2016 AGATGATACCCCTTTATGGCAGAAAGTGAAGAGAACTAAAAAGCCTCTTGTAGAAAGT 2075
Db 9877 AACTGATATCATATAATGGCAGAAAGTGAAGAGAACTAAAAAGTCTCTTGAAGAGAT 9936
QY 2076 GAAAGAGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAAAACGAAGATCATGGC 2135
Db 9937 GAAAAAGGAGAGTGAAGAAAACTGGCTTAAAGTCAACACTCAAAAAAATAAGATCATGGC 9996
QY 2136 ATCTGGTCCCATCACTTCATGGGAATAGATGGGGAACAGTGAACAGTGTGAGACTT 2195
Db 9997 ATCTGGTCCCATCACTTCATGGCAATAGATGTGGAANAAGTGAACAGGTCAGAGATTT 10056
QY 2196 TATTTTTTGGGGGCTCCAAATCACTG-CAGATGGTGAAGTGCAGCCATGAAATTTAAAG 2254
Db 10057 TATTTTCTT--GGGCTCCAGAAATTAAGTGTGGAATAGTGCAGCCATGAAATTTAAAG 10114
QY 2255 ACATTTACTCTTGAAGAAAGTTAAACCACTTAGATAGATATGAAAGAGAGAGACA 2314
Db 10115 ACATTTGTCTCTTGAATTTAAAGCTA--TGACAAACTTAGTGTATTTAAAGAGAGAGACA 10172
QY 2315 TTACCTTCCCAACAAAGCCCATCTAGTCAAGGCTATGGTTTTTCCAGTGGTCAATGATG 2374
Db 10173 TCATTTTCCGCAAAAGTCCATCTAATCAAGCTATGATTTTCCAGTAAATCATGATA 10232
QY 2375 GATGTGAGAGTTGGACTGTGAAGAAAGCTGAGCACTGGAAGTAATGATGCTTTTGAAGTGT 2434
Db 10233 GATGTGAGAGTTGGACCAATAAAGAGGCTGAATGCTAAAGAAATGATGCTTTTCAAACTGT 10292
QY 2435 GGTGTGTGAGAGAGTCTTGAAGTCCCTTGGAGTGAAGAGAGATCCAAACAGTCCATTC 2494
Db 10293 GGTGTGAGAGAGAGTCTTGAAGTCCCTTGGAGTGAAGAGATTAACCAAGTCAATTC 10352
QY 2495 TGAAGGAGATCAGCCCTGGGATTTCTTTGGAAGGAATGATCTAAAGCTGAAGTCCAGT 2554
Db 10353 TAAAGGAATCAACCCCTGAATATTCATTGGAAGAACTGATGCTGAATCTGAAGCTCCAAT 10412
QY 2555 ACTTTGGCCACTGATCAGAGAGTGTGACTGCTGGAAGAGACCTGATGCTGGAGGGA 2614
Db 10413 ACTTTGGCCACTGATGGAAGAGTGTGACTGCTGGAAGAGAGTCAATTTGGAAGAGA 10471
QY 2615 TTGGGGCAGGAGGAGAGGAGGAGACAGAGAGATGAGATGCTGGATGGATCACTGACT 2674
Db 10472 --CAGGACAGGAGGAAGAGGAGCAACAGAGAGATGA-----GATGGAATCATCAACT 10521
QY 2675 CGATGGACGTGAGTCTGGGTGAAGTCTCTGGAGTTGGTGAAGGACAGGAGGCGCTGCTG 2734
Db 10522 TAATGGACATGACTTTGAGCAAACTCCGGGAGATGAATGATGGACAGGAGAACTGGCATG 10581
QY 2735 CGCGATTATGCGGTGCAAGAGTGTGGACACCACTGAGCAACTGAAC----- 2783
Db 10582 CTGCAATTTATGGGTCTCAAGAGATGACATGACTTAGCACTGAACACACAAATGC 10641
QY 2784 ---TGAACGTAGTGTACTGAAACCTTAGTAGTTTATATTAATCTAGAAAAATAGTAATTC 2840
Db 10642 CAAGAACCAAGACAGAGACCTAAGTTTGAATGTGTATGATATAAATATATATATATAT 10701
QY 2841 ATATGATTTCAAAATTTTTCATATGTTGGTTAAGATATAAGATTTTCAATTTGATTT 2900
Db 10702 ATATATATATGATTTTTCCTATTTCTCAAAATCAATATTTGCTTTTATGACTGCTTT 10761
QY 2901 TTATCTTTGATTTTCTCTACTTATTTAAATTTGGGATTTTAACTATTTCTTCAATGACT 2960
Db 10762 GTTTCAGTTAGTGAATGTTCTTTAAATTTTATCATGTTGAGCAAAATTTTACTTCTCTT 10821
QY 2961 TGTATTTCTAATATTTACTTATTTCTTATTTTAAATGACATTTATTTATGATTTT 3020
Db 10822 TTAAGACTGAATAATACATCTCAATGAATAT--ACTACATTTTGTCTTCAATTTGTCTAT 10879
QY 3021 TCTAATAAAATCCAGTCTCTCTTTTAAAGAGACTTTAAATATTAATTTCTCTTTA 3080
```

```
Db 10880 GATAGCACATAGATTGCTTTAACTTTAGCTGTATGATTAATAATTTGATATGATCATGG 10939
QY 3081 GTGTTTTTACCAGTTCTTTTCCAGGCTACTCTTTTGGATTTTATTTGGTCTCTATCTTTCTCAA 3140
Db 10940 ATGTACAA--ATATCTTCAAGACCTTGTCTTTTAATCTTTTGTGTATACCAAGAGT 10997
QY 3141 GTTTTGAATGGCTAGTAACTCATTTATCTTTATTTTGTAAATAGCTCTTTAAAT 3200
Db 10998 GGAATTTACTGAATTAATTTGATTTTAAATTTTAAATTTTGGAGAACTGGCACTGAATG 11057
QY 3201 CATTTATCTTTGATACACAGCTTCAGTTCTPATGGCTTTAATAAAGTTTTTTTTTTTTTT 3260
Db 11058 TGTCAATTTACATTTCCCAACAAATGTTTCACTGGGT---TCCAGTTTCTACATCTCTG 11114
QY 3261 TTTTAAAGAAATGTCAATCTTTTGTGAAGTTTGAACAATGCTTTGAGCAATTAATTTAGGAT 3320
Db 11115 CCACCAATTTGTTTCCATGTTTTTAAAGAAAGTAGCCATTTCTAATGCATATGATGATAAC 11174
QY 3321 ATTTTGAATGGCT--CATGAGTATGCTTTTGTACTTGGCATTATTTGAAGTTT--ATGA 3376
Db 11175 CTCATTTGATGTTTGAAGATTCGCAATCTCTAGTGGTGGCATTGATTAATTTTCAATG 11234
QY 3377 TTTATGAATTTATGATGCTTTTTTTTGGGCATTAAGTCTATGGCATATTTTTTTGTGCT 3436
Db 11235 TTTATGACTATTTATGATATCTTTTCTGAGAAATGCTATTCAAGTCTTTTGTGTTTTT 11294
QY 3437 ATATCTTTAAATTAATAATTTGGCTT 3462
Db 11295 AAATGGATTTATTTGTTTTTTTTT 11320
```

RESULT 13

```
ABV72460
ID ABV72460 standard; DNA; 6620 BP.
XX
AC ABV72460;
XX
DT 29-JAN-2003 (first entry)
XX
DE A nucleotide sequence 3' of the bovine SRY gene.
XX
KW Sex selection; mammal; transgene; SRY gene; ss.
XX
OS Bos sp.
XX
PN WO200277637-A1.
XX
PD 03-OCT-2002.
XX
PF 22-MAR-2002; 2002WO-US008933.
XX
PR 22-MAR-2001; 2001US-0278155P.
XX
PA (INFI-) INFGEN INC.
XX
PI Forsberg EU, Eilertsen KJ, Bishop MD, Zheng Y, Leno GH;
XX
DR WPI; 2003-029952/02.
XX
PS Example 1; Page 46-48; 67pp; English.
```

CC The specification describes a method for pre-selecting the sex of mammalian offspring. The method permits the enrichment of X- or Y-chromosome bearing sperm in semen, by introducing a transgene on a sex chromosome. Expression of the transgene is operably linked to a promoter region that confers haploid-specific expression to the transgene. The mammal is useful in methods for producing a population of mammalian haploid cells that is enriched for haploid cells containing a specific


```
CC sex chromosome. The mammal, as well as the method is useful for pre-
CC selecting the sex of mammalian offspring. The present sequence represents
CC a sequence 3' of the bovine SKY gene, which was used as the homologous
CC arm in targeting vectors of the invention
XX
SQ Sequence 6620 BP; 2279 A; 1161 C; 1449 G; 1731 T; 0 U; 0 Other;
Query Match 16.5%; Score 589.4; DB 1; Length 6620;
Best Local Similarity 87.8%; Pred. No. 5.8; Indels 24; Gaps 7;
Matches 71; Conservative 0; Mismatches 76;
QY 1589 GAACATCCAGTGTTCAGAGTGGTTTGTAGAAAAGTCAGAGAACCCAGAGACCAAAATGGC 1648
Db |||||
XX 5818 GAAATTCCTGATGTGAAGCTGGTTTGTAGAAAAGTCAGAGAACCCAGAGATCAAAATGGC 5877
QY 1649 AACATCTCTGTATCATGGAAGCAAGAGAGTTCAGAAAACATCTATTTCTGCTTT 1708
Db |||||
XX 5878 AACATCTGTGGATCATGGAAGCAAGAGAGTTCAGAAAATATTTATTTCTGCTTT 5937
QY 1709 ATTGACTATGCAAAAGCCTTTGACTGTGGGGTGCACATAAACTGTGCAAAATTTCTGAA 1768
Db |||||
XX 5938 ATTGCTATGGAAGCAATGACTGTGTGATCACAGTACACTGTGGAATTTCTGAA 5997
QY 1769 GGGATGGGAATACAGAACCCACTGACTGTGTAATTTGTATGAGTCAAGAA 1828
Db |||||
XX 5998 CAGATGGGAATACAGAACCCACTTGACAGGCTCTTGAGAACTCTGTATGAGGTCAAGAA 6057
QY 1829 GCAACAGTTAGAACTGACACATGGAACACAGCTGTTTCAAGTGTAGAAAAGAGATGT 1888
Db |||||
XX 6058 GTAACAGTTAGAACTGACATGGAACATAGACTGTTCATATGGAAGAAAGATACAC 6117
QY 1889 CAAGCTGTATTTGTCAACCGGCTTGTTTAAC-----TTCTATGACAGAG-ACATCAT 1940
Db |||||
XX 6118 CAAGCGGCATATTTGTCAACCTGCTTATTTACCTGCGAGAGTACATGACAGATACATCAT 6177
QY 1941 GAGAAAGCTGGCTGGAAGAGACACAGCTGGAATCAAGTTCCTGGGAGAAATACAA 2000
Db |||||
XX 6178 GAGAAATGCTGGACTGGAAGAAACACAGCTGGAAATCAAGATTGCGAGGAAATATCAA 6237
QY 2001 TAACCTCAGATATGACATATACACCCCTTATGGCAGAAAGTGAAGGAACTAAAG 2060
Db |||||
XX 6238 TAACCTCTGATATGACATGACACCCCTTATGGCAGAAAGTGAAGGAACTAAAG 6297
QY 2061 CCTCTTGAAGTGAAGAGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAA 2120
Db |||||
XX 6298 CCGCTTAAAGAAAGTGAAGTGAAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAA 6357
QY 2121 AACGAAGATCATGCTGCTGCCATCCTTCAATGGAATATAGATGGGAAACAGTGA 2180
Db |||||
XX 6358 AACGAAGATCATGCTGCTGCCATCCTTCAATGGAATATAGATGGGAAAGATG-----GGA 6407
QY 2181 AACAGTGTACATTTATTTTGGGGGCTCCAAAATCACTGAGATGTGTGACTGCAGC 2240
Db |||||
XX 6408 AACAGTGTACAGATTTA--TTTGTGGCTTCCAAAATCACTGAGATGTGTGAGTGTGC 6465
QY 2241 CATGAATTAAGACACACTTACTCTCTTGGAGAAAGTTA--ACCACTTATAGATAGATA 2298
Db |||||
XX 6466 CATGAATTAAG--CACTTACTCTCTTGGAGAAAGTTATGACCACTTATAGATAGATA 6524
QY 2299 TTGAAAAGCAGACATTTACCTTCCAAACAAAGCCCATCTAGTCAAGGCTATGGTTTTT 2358
Db |||||
XX 6525 TTCAAAA--CAGAAACATTTACTTTGCCAACAAAGGCTCGCTAGTCAAGGCTATGGTTTTT 6583
QY 2359 CCAGTGTGATGATGATGAGAGTTGGACTGTGA 2395
Db |||||
XX 6584 CCTGTGTGATTTTGGATGTGAGAGTTGGACTGTGA 6620
RESULT 14
ABK14793
ID ABK14793 standard; DNA; 23653' BP.
XX
AC ABK14793;
```

```
XX
DT
XX
DE
XX
XX
KW Platelet-derived growth factor expression cassette insert used in pBC734.
KW Platelet-derived growth factor; PDGF; transgenic; milk; gene;
KW wound healing; diabetic foot ulcer; decubitus ulcer; ophthalmic disease;
KW venous stasis ulcer; periodontal regeneration; bone formation;
XX prosthetic vascular graft; pBC734; ds.
OS Homo sapiens.
XX
PN WO200198520-A1.
XX
PD 27-DEC-2001.
XX
XX 19-JUN-2001; 2001WO-US041044.
XX
XX 19-JUN-2000; 2000US-0212406P.
XX
XX (GENZ ) GENZYME TRANSGENICS CORP.
XX
XX Echelard Y, Meade H, Eichner W, Sommermeyer X;
XX
XX WPI; 2002-083329/11.
XX
XX Production of platelet derived growth factor (PDGF) comprises expression
XX in the milk of a non-human transgenic animal.
XX
XX Example 1; Fig 1; 59pp; English.
XX
CC The invention relates to the production of platelet-derived growth factor
CC (PDGF) comprising a transgenic mammal whose somatic and germ cells
CC comprise a nucleic acid sequence (I) encoding PDGF, operably linked to a
CC promoter directing expression into mammalian gland epithelial cells, and
CC obtaining the milk from the transgenic mammal where at least 30% of the
CC PDGF in the milk is as a dimer. Also described is a method of producing a
CC transgenic mammal capable of expressing an active PDGF molecule in milk.
CC Pharmaceutical compositions can be obtained from this milk and can be
CC used to stimulate or enhance the wound healing process, in particular
CC diabetic foot ulcers, decubitus ulcers and venous stasis ulcers.
CC Transgenic PDGF (II) can also be used in the treatment of periodontal
CC regeneration, stimulation of bone formation, ophthalmic diseases or
CC healing of prosthetic vascular grafts. (II) can also be used for non-
CC medical applications, e.g., as a supplement for cell culture media or as
CC a component of diagnostic kits. The present sequence represents the
CC platelet-derived growth factor expression cassette insert used in pBC734
XX vector of the invention
XX
SQ Sequence 23653 BP; 6803 A; 5184 C; 5014 G; 6650 T; 0 U; 2 Other;
```

```
Query Match 15.6%; Score 557.9; DB 1; Length 23653;
Best Local Similarity 78.4%; Pred. No. 2.1;
Matches 820; Conservative 0; Mismatches 191; Indels 35; Gaps 13;
QY 1744 CAATTAACCTGTGGAATAATCTGAAAGGGATGGGAATACCAGACCACCTGACCTGACTCT 1803
Db |||||
XX 2456 CACACCAACTGGAATAATTTCTTCAAGAG--AGAAATACCAGACCACCTTACTTGTCTCT 2513
QY 1804 GAAAATTTGTATGAGGTGAGGAAGCAACAGTTAGAACTGGACATGGAAACAACAGACTG 1863
Db |||||
XX 2514 GAGAAATCTGTTTGTCTGCTCA--GAAGCAACAGTTAGAACACAGATGGAACAACAGACTG 2572
QY 1864 GTTCCAGTAGGAAGAGAGTATGTCAAGGCTGTATATGTTCACCGGCTTGTAACTT 1923
Db |||||
XX 2573 GTTCCAAATCAGGAAGAGAGTATGTCAAGGCTGTATATGTTCACCGCTGATTTAACTT 2632
QY 1924 CTATGCAGAG-ACATCATGAGAAACCTGGGCTGGAAGAAAGCAACAGCTGGGAATCAAGAT 1982
Db |||||
XX 2633 ATATGCATAGTACATAATACAAATGCCAGGCTGGATGATCGCAAGCTGGGAATCAAGAT 2692
QY 1983 TGCCGGGAGAAATAGCAATAAATCACTGATATGCAATGATACCACTTATGCGAGAAG 2042
Db |||||
```


Db 1056 CCGTCTAGTCAAGGCTATGCTTTTCCAGTAGTCAATGATGGAGTGCAGAGTTGGACTGT 997
 Qy 2394 GAAGAAAGCTGAGCACTGAAGAAATGATGCTTTTGAAGTGGTGGTGGAGAACTCTT 2453
 Db 996 GAAGAAAGCTGAGGCTGAAGAAATGATGCTTTTGAAGTGGTGGTGGAGAACTCTT 997
 Qy 2454 GAGAGTCCCTTGGAGTCAAGAGATCCAAACAGTCCATTCTGAGAGATCAGCCCTGG 2513
 Db 936 GAGAGTCCCTTGGAGTCAAGAGAAATCCAAACAGTCCATTCTGAGAGATCAGCCCTGG 877
 Qy 2514 CATTTCTTTTGAAGGAATGATGCTTAAAGCTGAAACTCCAGTACTTTTGGCCACCTGATCAG 2573
 Db 876 GATTTCCTTTGGAAGGAATGATGCTTAAAGCTGAAACTCCAGTACTTTTGGCCACCTCAGCA 817
 Qy 2574 AAGAGTCACTACCTGGAAGAAAGACCCCTGATGCTGGAGGGATTTGGGGGAGGAGGAAG 2633
 Db 816 AAGAGTTTGAACCTATTGGAAGAAAGACTCTGATGCTGGAGGAACCTGGGGGAGGAGGAAG 757
 Qy 2634 GGGACGACAGAGGATGAGATGGCTGGATGGATGCGATCACTGATCGATGACGTGAGTCTGGG 2693
 Db 756 GGGACACAGAGGATGAGATGGCTGGAGCGGATCACCCTACTCGATGACGTGAGTCTGAG 697
 Qy 2694 TGAACCTCTGGAGTTGGTGTGATGGAACAGGAGGCTGCTGCGCGGATTCATGGGFTAC 2753
 Db 696 TGAACCTCTGGAGTTGGTGTGATGGAACAGGAGGCTGCGATGCTCGGATTCATGGGFTGC 637
 Qy 2754 AAAGAGTTGGACAGACTGACCACTGAACCTGAACCTGAACCTGAACCTTAGT 2813
 Db 636 AAAGAGTCAACAGACTGACCGACTGAACCTGAACCTGAACCTGAACCTGTTACAGA 577
 Qy 2814 TTATATTACTCAGAAATAGTAAATTCATATGATGATTTCAAAATTAATTTTCA 2864
 Db 576 AATTACAGGAAATATGATCTTTGTTAAACATTTGCAATTTTGTAAA 526

RESULT 18

AAD26492/c
 ID AAD26492 standard; DNA; 1813 BP.
 AC AAD26492;
 DT 26-MAR-2002 (first entry)
 DE Sheep alpha (1, 3) galactosyltransferase post-exon 4 DNA #3.
 KW Sheep; xenotransplantation; Gal-alpha (1,3)Gal; GAL determinant; surgery;
 KW alpha(1,3) galactosyltransferase; alpha (1, 3) GT; drug screening;
 KW gene therapy; ds.
 XX Ovis sp.
 OS
 XX WO200180896-A2.
 PN
 XX 22-NOV-2001.
 PD
 XX 14-MAY-2001; 2001WO-US015765.
 PF
 XX 15-MAY-2000; 2000US-0204148P.
 PR
 XX 13-JUN-2000; 2000US-00593316.
 XX
 PA (GERO-) GERON CORP.
 XX
 XX Denning C, Clark J;
 PI
 XX WPI; 2002-089848/12.
 DR
 XX New ovine tissue devoid of antibody-detectable Gal-alpha(1,3)Gal
 PT determinants, for xenotransplantation, and in the treatment of the human
 PT body by surgery or therapy.
 XX
 XX Claim 18; Page 79; 86pp; English.
 PS
 XX The patent discloses immunologically compatible animal tissue, suitable
 CC

CC for xenotransplantation into human patients. The invention also relates
 CC to an ovine tissue devoid of antibody-detectable Gal-alpha (1,3)Gal (GAL)
 CC determinants which are made by alpha(1,3) galactosyltransferase (GT). The
 CC ovine tissue is useful for treatment of human body by surgery or therapy
 CC and in xenotransplantation, by transplanting the ovine tissue into a
 CC mammal having circulating antibody against Gal alpha(1,3)GT determinants.
 CC Polynucleotide constructs of the invention are useful for inactivating an
 CC alpha(1,3) GT gene in an ovine cell. Alpha (1,3) GT sequences are useful
 CC for drug screening and for the production of GAL containing synthetic
 CC oligosaccharides. Sequences of the invention are also useful in gene
 CC therapy. The present sequence is alpha (1,3) GT post-exon 4 DNA
 XX
 SQ Sequence 1813 BP; 477 A; 391 C; 428 G; 515 T; 0 U; 2 Other;
 Query Match 10.9%; Score 388.9; DB 1; Length 1813;
 Best Local Similarity 88.2%; Pred. No. 88;
 Matches 455; Conservative 0; Mismatches 56; Indels 5; Gaps 3;
 Qy 2282 CCAACTAGATAGCATATTGAAAGCAGAGACATTACCTTGCACAAAGCCCATCTA- 2340
 Db 936 CTAAATGTTTATAGCATATTCAAAAGCAGAGACATTACTTTGCCACAAAGGTCATCTAT 877
 Qy 2341 GTCAAGGCTATGTTTTCAGTGGTCAATGATGATGAGTGGACTGTGAAGAAA 2400
 Db 876 GTCAAGGCTATGTTTTCAGTGGTCAATGATGATGAGTGGACTGTGAAGAA 817
 Qy 2401 GGTGAGCACTGAAGATTTGCTTTTGAAGTGGTGGTGGAGAGACTCTTGAGAGTC 2460
 Db 816 GGTGAGCCCCGAGAAATCGATGCTTTTGAAGTGGTGGTGGAGAGACTCTTGAGAGTC 757
 Qy 2461 CTTTGGACTGCAAGGAGATCCAAACAGTCCATTCTTGAAGAGAGATCAGCCCTGGATTCT 2520
 Db 756 CTTTGGACTGCAAGGAGATCCAAACAGTCCATTCTTGAAGAGATCAGCCCTGGATTCT 697
 Qy 2521 TTGGAAGATGATGCTAAAGCTGAACTCAGTACCTTTGGCCACCTGATCAGAGAGCT 2580
 Db 696 TTGGAAGATGATGCTAAAGCTGAACTCAGTACCTTAAGCCACCTCATGAGAGAGT 637
 Qy 2581 ---GACTCACTGGAAGAACCTGATGCTGGAGGGATTGGGGGAGAGAGAGGGGA 2637
 Db 636 TGACTCATATGGAAGAGACTCATGAGCTGGAGGGATTGGGGGAGAGAGAGGGGA 577
 Qy 2638 CGACAGAGATGAGATGGCTGGATGGATCATGCTGATGATGAGTGGTGGTGGTGA 2697
 Db 576 CGACAGAGATGAGATGGCTGGATGGATCAACCAACTCAATGAGACAGAGTTTGGTGA 517
 Qy 2698 CTCTGGAGTTGGTGGATGAGAGAGGAGGCTGCTCTGGGGGATT-CATGGGGTCACAAA 2756
 Db 516 CTCTAGGAGTTGGTGGATGAGAGAGGAGGCTGGGCTGCTGCATTCATGGGGTCGAAA 457
 Qy 2757 GAGTTGGACAGACTGAGCAACTGAACCTGAACCTGAA 2792
 Db 456 GAGTGGACAGACTGAACGACTGAACCTGAACCTGAA 421

RESULT 19

ADE81111
 ID ADE81111 standard; DNA; 967 BP.
 XX
 AC ADE81111;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Bovine 1-8U protein-encoding gene upstream region.
 XX
 KW bovine 1-8U protein; bovine Leu-13 protein; interferon; bovine pregnancy;
 KW detecting pregnancy; ds.
 XX
 OS Bos taurus.
 XX
 PN US2003143601-A1.
 XX 31-JUL-2003.
 PD

XX Bovine EST associated with lactation/muscle/fat deposition #6161.
DE Bovine, ss; EST; expressed sequence tag; lactation; LMFD;
XX muscle deposition; fat deposition; genome mapping; gene identification;
KW gene analysis; cattle breeding.
XX Bos Taurus.
XX US2002137139-A1.
XX 26-SEP-2002.
XX 24-SEP-2001; 2001US-00960352.
XX 12-JAN-1999; 99US-0115707P.
XX 11-JAN-2000; 2000US-00480902.
XX (BYAT/) BYATT J C.
XX (MATH/) MATHIALAGAN N.
XX (TAON/) TAO N.
XX (WARR/) WARREN W C.
XX Byatt JC, Mathialagan N, Tao N, Warren WC;
XX WPI; 2003-110599/10.
XX New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and analysis,
PT cattle breeding, or for genetically improving cattle.
XX Claim 2; SEQ ID NO 6161; 245pp; English.
XX The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMFD), derived from
CC cattle, and the LMFD nucleic acid can specifically hybridize to a second
CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
CC appearing as ABX34836-ABX49947, or complements of them. Also included are
CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
CC acid linked to a promoter and a 3' non-translated sequence that
CC functions in the cell to cause termination of transcription and addition
CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
CC (2) determining a level or pattern of a molecule in a bovine cell or
CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
CC of the 15112 nucleic acid sequences or its complement or fragment) with a
CC complementary nucleic acid molecule obtained from the bovine cell or
CC tissue, where hybridisation between the marker nucleic acid and the
CC complementary nucleic acid permits the detection of the molecule; and (b)
CC detecting the level or pattern of the complementary nucleic acid, where
CC the detection of the complementary nucleic acid is predictive of the
CC level or pattern of the molecule. The LMFD nucleic acid is used for
CC determining a level or pattern of a molecule in a bovine cell or tissue.
CC It is useful for genome mapping, gene identification and analysis, cattle
CC breeding, preparation of constructs for use in cattle gene expression, or
CC for genetically improving cattle. The present sequence is one of the
CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
CC present sequence was not shown in the specification but was obtained in
CC electronic format from the USFTO web site:
CC seqdata.uspto.gov/sequence.html?DocID=20020137139

XX Sequence 436 BP; 156 A; 84 C; 96 G; 100 T; 0 U; 0 Other;
Query Match 10.2%; Score 363.2; DB 1; Length 436;
Best Local Similarity 93.4%; Pred. No. 4.3e+02;
Matches 410; Conservative 0; Mismatches 23; Indels 6; Gaps 3;
XX 1892 GCGTGTATTGTGTCACCGGGCTGTTTAACTTCTATGTCAGAG-ACATCATGAGAAAGCT 1950
DB 1 GGCAAGTATTGTGTCACCGCTGCTTATTTAACTTCTATGTCAGAGTACATCATGAGAAAGCT 60
XX 1951 GGGCTGGAGAGACACAGCTGGATTCAGATTCCCGGGGAGAAATAGCATATACCTCAGA 2010
DB 61 GGAGTGGAGAAACACAGCTGGATTCAGATTCCCGGGGAGAAATATCAATAAAGCTCAGA 120

QY 2011 TATGCAGATGATACACCCCTTATGGCAGAAAGTGAAGAGAACTTAAAAAGCCTCTTGATG 2070
DB 121 TATGCAGATGATACACCCCTTATAGCAGAAAGTGAAGAGAACTTAAAAAGCCTCTTGATG 180
QY 2071 AAGGTGAAAGAGAGAGATGAAAAAGTTGGCTTAAAGCTCAACATTCAGAAAAACGAGATC 2130
DB 181 AAAGTGAATTTGGAGAGTGAAGAAAGTTGGCTTAAAGTTCAACATTCAGAAAAATGAAGATC 240
QY 2131 ATGGCATCTGCTCCCATCTTCACTTCAATGGAATAGATGGGAAAACAGTGGAAACAGTGTCA 2190
DB 241 ATGGCATCTGCTCCCATCTTCACTTCAATGGAATAGATGGGAAAACAGTGGAAACAGTGTCA 300
QY 2191 CACTTTATTTTTGGGGGGCTCCAAAATCACTGCAGATGGTGAAGTGGAGCCATGAATTA 2250
DB 301 CACTTTATTTCTTT---GGGCTCCAAAATCACTGCAGATGGTGAAGTGGAGCCATGAATTA 357
QY 2251 AAAGACACTTACTCTTGGAGAGAAAGTTA--ACCAACCTAGATAGCATATTGAAAAAGCA 2308
DB 358 AAAGACGCTTACTCTTGGAGAGAAAGTTATGACCAACCTAGATAGCATATTGAAAAAGCA 417
QY 2309 GAGACATTACCTTGCCAAAC 2327
DB 418 GAGACATTACTTTGCCAAAC 436
RESULT 23
ABX39925
ID ABX39925 standard; cDNA; 432 BP.
XX AC ABX39925;
XX DT 20-FEB-2003 (first entry)
XX DE Bovine EST associated with lactation/muscle/fat deposition #5090.
XX KW Bovine, ss; EST; expressed sequence tag; lactation; LMFD;
KW muscle deposition; fat deposition; genome mapping; gene identification;
KW Gene analysis; cattle breeding.
XX OS Bos Taurus.
XX PN US2002137139-A1.
XX PD 26-SEP-2002.
XX PF 24-SEP-2001; 2001US-00960352.
XX PR 12-JAN-1999; 99US-0115707P.
XX PR 11-JAN-2000; 2000US-00480902.
XX (BYAT/) BYATT J C.
XX (MATH/) MATHIALAGAN N.
XX (TAON/) TAO N.
XX (WARR/) WARREN W C.
XX Byatt JC, Mathialagan N, Tao N, Warren WC;
XX WPI; 2003-110599/10.
XX New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and analysis,
PT cattle breeding, or for genetically improving cattle.
XX Claim 2; SEQ ID NO 5090; 245pp; English.
XX The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMFD), derived from
CC cattle, and the LMFD nucleic acid can specifically hybridize to a second
CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
CC appearing as ABX34836-ABX49947, or complements of them. Also included are
CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
CC acid linked to a promoter and a 3' non-translated sequence that
CC acid linked to a promoter and a 3' non-translated sequence that

CC functions in the cell to cause termination of transcription and addition
 CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
 CC (2) determining a level or pattern of a molecule in a bovine cell or
 CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
 CC of the 15112 nucleic acid sequences or its complement or fragment) with a
 CC complementary nucleic acid molecule obtained from the bovine cell or
 CC tissue, where hybridisation between the marker nucleic acid and the
 CC complementary nucleic acid permits the detection of the molecule; and (b)
 CC detecting the level or pattern of the complementary nucleic acid, where
 CC the detection of the complementary nucleic acid is predictive of the
 CC level or pattern of the molecule. The LMFD nucleic acid is used for
 CC determining a level or pattern of a molecule in a bovine cell or tissue.
 CC It is useful for genome mapping, gene identification and analysis, cattle
 CC breeding, preparation of constructs for use in cattle gene expression, or
 CC for genetically improving cattle. The present sequence is one of the
 CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
 CC present sequence was not shown in the specification but was obtained in
 CC electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?DocID=20020137139
 XX
 CC Sequence 432 BP; 146 A; 87 C; 100 G; 99 T; 0 U; 0 Other;

Query Match 10.1%; Score 360.3; DB 1; Length 432;
 Best Local Similarity 94.3%; Pred. No. 4.4e+02;
 Matches 384; Conservative 0; Mismatches 22; Indels 1; Gaps 1;
 Qy 1589 GAACCTCCAGTTGTTCAAGCTGGTTTGTAGAAAAGTTCAGAGAACCCAGACCAATGGCC 1648
 Db 26 GATCTCCAGTTGTTCAAGCTGGTTTGTAGAAAAGTTCAGAGAACCCAGACCAATGGCC 85
 Qy 1649 AACATCTCTGTATCATGGAAAAGCAAGAGAGTTCAGAAAACATCTATTCTGCTTT 1708
 Db 86 AACATCCGCTGAATCATCGAAAAGCAAGAGAGTTCAGAAAACATCTATTCTGCTTT 145
 Qy 1709 ATTGACTATGCAAAAGCCTTTGACTGTGGGGTTCACAAATAAATCTGGAAAATTCGAAA 1768
 Db 146 ATTGACTATGCCAGGCTTTGACTGTGTGATCACAATAAATCTGGAAAATTCGAAA 205
 Qy 1769 GGGATGGGAATACAGACCACTGACCTGACTCTTTGAAAAATTTGTATGACGTCAAGAA 1828
 Db 206 AAGATGGGAATACAGACCACTGACCTGACTCTTTGAAAAATTCGATGACGACCAAGAA 265
 Qy 1829 GCACAGCTTAGAATGACATGGAACACACAGAGTGGTCCAGTAGGAAAGAGTATCT 1888
 Db 266 GCACAGCTTAGAATGACATGGAACACACAGAGTGGTCCAGTAGGAAAGAGTACGT 325
 Qy 1889 CAAGGCTGTATATGTGACCCGGCTTGTATTAATCTTATGACAG-ACATCATGAGAAAC 1947
 Db 326 CAAGGCTGTATATGTGACCCGGCTTGTATTAATCTTATGACAGTACATTAAGAAAC 385
 Qy 1948 CTGGGCTGGAAGNAGCAGAGCTGGAAATCAAGATTGCCGGGAGAA 1994
 Db 386 TCTGGGCTGGAAGNAGCAGAGCTGGAAATCAAGATTGCCGGGAGAA 432

RESULT 24
 ID ABX39519
 XX ABX39519 standard; cDNA; 433 BP.
 XX
 AC ABX39519;
 XX
 DT 20-FEB-2003 (first entry)
 XX
 DE Bovine EST associated with lactation/muscle/fat deposition #4684.
 XX
 KW Bovine, ss; EST; expressed sequence tag; lactation; LMFD;
 KW muscle deposition; fat deposition; genome mapping; gene identification;
 KW gene analysis; cattle breeding.
 XX
 OS Bos Taurus.
 XX
 PN US2002137139-A1.
 XX

PD 26-SEP-2002.
 XX
 PF 24-SEP-2001; 2001US-00960352.
 XX
 PR 12-JAN-1999; 99US-0115707P.
 PR 11-JAN-2000; 2000US-00480902.
 XX
 PA (BYAT/) BYATT J C.
 PA (MATH/) MATHIALAGAN N.
 PA (TAON/) TAO N.
 PA (WARR/) WARREN W C.
 XX
 BYatt JC, Mathialagan N, Tao N, Warren WC;
 XX
 WPI; 2003-110599/10.
 DR
 XX New nucleic acid associated with lactation, and muscle and fat
 PT deposition, useful for genome mapping, gene identification and analysis,
 PT cattle breeding, or for genetically improving cattle.
 XX
 PS Claim 2; SEQ ID NO 4684; 245pp; English.
 XX
 CC The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMFD), derived from
 CC cattle, and the LMFD nucleic acid can specifically hybridise to a second
 CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
 CC appearing as ABX34836-ABX49947, or complements of them. Also included are
 CC: (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
 CC acid linked to a promoter and a 3' non-translated sequence that
 CC functions in the cell to cause termination of transcription and addition
 CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
 CC (2) determining a level or pattern of a molecule in a bovine cell or
 CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
 CC of the 15112 nucleic acid sequences or its complement or fragment) with a
 CC complementary nucleic acid molecule obtained from the bovine cell or
 CC tissue, where hybridisation between the marker nucleic acid and the
 CC complementary nucleic acid permits the detection of the molecule; and (b)
 CC detecting the level or pattern of the complementary nucleic acid, where
 CC the detection of the complementary nucleic acid is predictive of the
 CC level or pattern of the molecule. The LMFD nucleic acid is used for
 CC determining a level or pattern of a molecule in a bovine cell or tissue.
 CC It is useful for genome mapping, gene identification and analysis, cattle
 CC breeding, preparation of constructs for use in cattle gene expression, or
 CC for genetically improving cattle. The present sequence is one of the
 CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
 CC present sequence was not shown in the specification but was obtained in
 CC electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?DocID=20020137139
 XX
 SQ Sequence 433 BP; 153 A; 74 C; 105 G; 101 T; 0 U; 0 Other;

Query Match 10.0%; Score 357.1; DB 1; Length 433;
 Best Local Similarity 90.5%; Pred. No. 4.5e+02;
 Matches 401; Conservative 0; Mismatches 29; Indels 13; Gaps 2;
 Qy 1943 GAACGCTGGCTGGAGNAGCAGAGCTGGAAATCAAGATTGCCGGGAGAAATAGCAATA 2002
 Db 1 GAACGCTGGCTGGAGNAGCAGAGCTGGAAATCAAGATTGCCGGGAGAAATATCAATA 60
 Qy 2003 ACCTCAGATATGCAGATGATACCAACCTTTATGGCAGAAAGTGAAGAGAGAACTAAAAAGCC 2062
 Db 61 ACCTCAGATATGCAGATGATACCTACCTTTATGGCAGAAAGTGAAGAGAGAACTAAAGAGCC 120
 Qy 2063 TCTTGATCAAGTCAAAAGAGAGAGTGAAGAAAGTGGCTTAAAGCTCAACATTCAGAAA 2122
 Db 121 TCTTGATCAAGTCAAAAGAGAGAGTGAAGAAAGTGGCTTAAAGCTCAACATTCAGAAA 180
 Qy 2123 CGAAGATCATGGCATCTGGTCCCATCACTTCATGGGAAATAGATGGGAAACAGTGGAA 2182
 Db 181 CTAAGATCATGGCATCTGGTCCCATCACTTCATGGGAAATAGATGGGAAACAGTGG-- 237
 Qy 2183 CAGTGTGACACTTTATTTTGGGGGCTCCAAATCACTGCAGATGCTGACTGCACCA 2242

238 -----CTGACTTTTATTTGGGGGCTCCAAAATCACTGCAGATGGTAACCTGAGGCA 289
 2243 TGAATTTAAAGACACTTACTCTTGGAAAGAAAGTTA--ACCAACCTAGATAGCATATT 2300
 290 TGAATTTAAAGACACTTACTCTTGGAAAGAAAGTTA--ACCAACCTAGATAGCATATT 349
 2301 GAAAACAGACAGATTAACCTTGGCAAAAGCCCATCTAGTCAGAGCTATGTTTTTCC 2360
 350 AAATAGCAGAGACATCACTTTGTCAACAAAGGTCCATCTAGTCAGAGCTATGTTTTTTC 409
 2361 AGTGGTCATGTGATGTGAGA 2383
 410 AGTGGTCATGTGATGTGAGA 432

RESULT 25
 ABX40182
 ID ABX40182 standard; cDNA; 433 BP.
 XX AC ABX40182;
 XX DT 20-FEB-2003 (first entry)
 XX DE Bovine EST associated with lactation/muscle/fat deposition #5347.
 XX KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
 XX KW muscle deposition; fat deposition; genome mapping; gene identification;
 XX KW gene analysis; cattle breeding.
 XX OS Bos Taurus.

US2002137139-A1.

26-SEP-2002.

24-SEP-2001; 2001US-00960352.

12-JAN-1999; 99US-0115707P.

11-JAN-2000; 2000US-00480902.

(BYAT/) BYATT J C.

(MATH/) MATHIALAGAN N.

(TAON/) TAO N.

(WARR/) WARREN W C.

Byatt JC, Mathialagan N, Tao N, Warren WC;

WPI; 2003-110599/10.

New nucleic acid associated with lactation, and muscle and fat

deposition, useful for genome mapping, gene identification and analysis,

cattle breeding, or for genetically improving cattle.

Claim 2; SEQ ID NO 5347; 245pp; English.

The invention relates to a purified nucleic acid molecule associated with lactation or muscle and fat deposition (designated LMFD), derived from cattle, and the LMFD nucleic acid can specifically hybridise to a second nucleic acid molecule comprising any of 15112 nucleotide sequences, appearing as ABX34836-ABX49947, or complements of them. Also included are: (1) a transformed cell having a nucleic acid comprising an LMFD nucleic acid linked to a promoter and a 3' non-translated sequence that functions in the cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and (2) determining a level or pattern of a molecule in a bovine cell or tissue comprising: (a) incubating a marker nucleic acid (comprising any of the 15112 nucleic acid sequences or its complement or fragment) with a complementary nucleic acid molecule obtained from the bovine cell or tissue, where hybridisation between the marker nucleic acid and the complementary nucleic acid permits the detection of the molecule; and (b) detecting the level or pattern of the complementary nucleic acid, where the detection of the complementary nucleic acid is predictive of the level or pattern of the molecule. The LMFD nucleic acid is used for

CC determining a level or pattern of a molecule in a bovine cell or tissue.
 CC It is useful for genome mapping, gene identification and analysis, cattle
 CC breeding, preparation of constructs for use in cattle gene expression, or
 CC for genetically improving cattle. The present sequence is one of the
 CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
 CC present sequence was not shown in the specification but was obtained in
 CC electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?DocID=20020137139

Sequence 433 BP; 159 A; 81 C; 94 G; 99 T; 0 U; 0 Other;
 Query Match 9.9%; Score 353.3; DB 1; Length 433;
 Best Local Similarity 90.0%; Pred. No. 4.6e+02;
 Matches 389; Conservative 0; Mismatches 42; Indels 1; Gaps 1;

QY 1701 TCTGCTTTTATGACTATGCATAAAGCCCTTGACTGTGGGGTCACATAAATGTTGAAAA 1760
 Db 2 TCTGCTTTTATGACTATGCATAAAGCCCTTGACTGTGGGGTCACATAAATGTTGAAAA 61
 QY 1761 TTCTGAAAGGATGGGAATACCAACACCTGACTCTTGAAGAAATTTGTATGCAG 1820
 Db 62 TTCTGAAAGATGGGAATACCTGCCACCTGACTCTTGAAGAAATTTGTATGCAG 121
 QY 1821 GTCAGGAAGCAACAGTTAGAACTGTGACATGGAACACAGACTGGTTCAGTAGAAAAAG 1880
 Db 122 GTCAGGAAGCAACAGTTAGAACTGTGACATGGAACACAGACTGGTTCAGTAGAAAAAG 181
 QY 1891 GAGTATGTCAGAGCTGTATATTGTCACCCGCTGTTTAACTTCTATGCAGAG-ACATCA 1939
 Db 182 GAGTACGTCAGAGCTGTATATTGTCACCCGCTGTTTAACTTCTATGCAGAGTACATCA 241
 QY 1940 TCAGAAAGCTGGGCTGGAAGAACACAAAGCTGGAATCAAGATGCCGGGAGAAATACGA 1999
 Db 242 TGAGAAATGCTGGACTGGATGAAGCCCAAGCTGGAGTCAAAATTCGCCGGGAGAGATATCA 301
 QY 2000 ATAACTCAGATATGCAGATATACCACTTATGGCAGAAAGTGAAGGAAGTGAAGGAACCTAAAA 2059
 Db 302 ATAACTCAGATATGCAGATATACCACTTATGGCAGAAAGTGAAGGAAGTGAAGGAACCTAAAA 361
 QY 2060 GCTCTTGTATGAAGCTGAAAGAGGAGAGTGAAGAAAGTTGGCTTTAAAGCTCAACATTCAGA 2119
 Db 362 GCTCTTGTATGAAGCTGAAAGAGGAGAGTGAAGAAAGTTGGCTTTAAAGCTTCAACATTCAGA 421
 QY 2120 AAACGAAGATCA 2131
 Db 422 AAACGAAGATCA 433

RESULT 26

AAH21025/c

ID AAH21025 standard; DNA; 1848 BP.

XX AC AAH21025;

XX DT 31-AUG-2001 (first entry)

XX DE Bovine-derived DNA fragment contig1 445_67.

XX KW Bovine; digital DNA signature; breeding; animal product origin;

XX KW identification; genetic association; breed; population; race; ds.

XX OS Bos taurus.

XX PN DE19959751-A1.

XX PD 13-JUN-2001.

XX PF 11-DEC-1999; 99DE-01059751.

XX PR 11-DEC-1999; 99DE-01059751.

XX PA (FRIE/) FRIES H R.

PA (DURS/) DURSTEWITZ G.

QY 2051 AACTAAAAACCTCTTGATGAAGTGAAGAGGAGTGAAGAGTGGCTTAAAGCTCA 2110
 Db 144 AACTAAAGAGCTCTTTGATGAAGTGAAGAGGAGTGAAGAGTGGCTTAAAGCTCA 203
 QY 2111 ACATTGAGAAACGAGATCATGCTGCTCCATCACTTCTGGAATAGATGGG 2170
 Db 204 ACATTGAGAAACGAGATCATGCTGCTCCATCACTTCTGGAATAGATGGG 263
 QY 2171 AAACAGTGAACAGTGTCAAGCTTTATTTTGGGGGCTCCAAATCACTGAGATGG 2230
 Db 264 AAACAGTGAACAGTGTCAAGCTTTATTTGGGGGCTCCAAATCACTGAGATGG 321
 QY 2231 TGACTGCAACCTGAATTAAGACACTTACTCTTGGAGAAAGTTA--ACACCT 2288
 Db 322 TGATTGCAACCTGAATTAAGACACTTACTCTTGGAGAAAGTTAAGCAACCT 381
 QY 2289 AGATAGCATATTGAAAGCAGAGACATTACCTTGCCAAAGCCCATCTAGTC 2343
 Db 382 AGATAGCATATTGAAAGCAGAGACATTACCTTGCCAAAGCTCCATCTAGTC 436

RESULT 28

ABQ75995
 ID ABQ75995 standard; DNA; 31412 BP.

XX
 AC ABQ75995;

XX
 DT 05-NOV-2002 (first entry)

XX
 XX Sheep PrP gene nucleic acid sequence.

XX
 DE Sheep; PrP; prion; transmissible spongiform encephalopathy; TSE; scrapie;
 KW breeding; selection; gene; ds.

XX
 XX Ovis aries.

XX
 PN GB2371048-A.

XX
 PD 17-JUL-2002.

XX
 XX 10-JAN-2001; 2001GB-00000702.

XX
 PR 10-JAN-2001; 2001GB-00000702.

XX
 PA (UYVO-) UNIV YORK.

XX
 PI Ross J, Bowles D;

XX
 DR WPI; 2002-577556/62.

XX
 PT Assay for comparing nucleic acid sequence with a reference, useful for
 PT determining prion gene alleles, comprises determining the profile of
 PT heteroduplexes.

XX
 PS Claim 2 (i) (a); Fig 1; 36pp; English.

XX
 CC The invention relates to an assay for determining allelic variations in
 CC prion protein genes, involving comparing a nucleic acid sequence with a
 CC reference. The method is used for determining the PrP allele in an animal
 CC (or human) in order to assess susceptibility to transmissible spongiform
 CC encephalopathies (TSE), most especially scrapie in sheep. This assists in
 CC the selection, for breeding, of animals having the alleles known to be
 CC associated with lowest risk. The method is reproducible and rapid,
 CC provides high resolution between allelic variants and is suitable for
 CC processing large numbers of samples. The current sequence represents a
 CC sheep PrP gene, that is used in the context of the invention as a
 CC comparison nucleic acid

XX
 SQ Sequence 31412 BP; 8889 A; 6517 C; 6464 G; 9542 T; 0 U; 0 Other;

Query Match 9.6%; Score 344.199; DB 1; Length 31412;

Best Local Similarity 69.7%; Pred. No. 6.9;

Matches 492; Conservative 0; Mismatches 208; Indels 6; Gaps 2;
 QY 2410 TGAAGAATTGATGCTTTTGAACCTGTGGTGTGGAGAGACCTCTTGAGAGTCCCTTGGACT 2469
 Db 17733 TGATTAACAGTAGTGTGTTGAACCTGTGGTGTGGAGAGACCTCTTGAGAGTCCCTTGGACT 17792
 QY 2470 GCAAGAGATCAACACAGTCCATCTCTGAAGAGATCAGCCCTGGATTTCTTTGGAAGGA 2529
 Db 17793 GCAAGAGATCAACACAGTCCATCTCTGAAGAGATCAGCCCTGGATTTCTTTGGAAGGA 17852
 QY 2530 ATGATGCTTAAAGCTGAAACCTCAGTACTTTGGCCACCTGATCAAGAGAGTCACTG 2589
 Db 17853 ATGATGCTTAAAGCTGAAATTCAGTACTTTGGCCACCTCATGCAAGAGTCACTG 17912
 QY 2590 GAAAGACCTGATCTCTGGAGGATTTGGGGGAGAGAGAGAGGACACAGAGATG 2649
 Db 17913 GAAAGACCTGATCTCTGGAGGATTTGGGGGAGAGAGAGAGGACACAGAGATG 17972
 QY 2650 AGATGGCTGGATGGCATCACTGACTTCGATGGACCTGAGTCTGGGTGAACCTCTGGAGTTG 2709
 Db 17973 AGATGGCTGGATGGCATCACTGACTTCGATGGACCTGAGTCTGGGTGAACCTCTGGAGTTG 18032
 QY 2710 GTGATGGACAGGAGGCTCTCTGGGGCGATTCATGGGGTCAAAAGAGTTGGACAGGA 2769
 Db 18033 TTGATGGACAGGAGGCTCTGGGGCGATTCATGGGGTCAAAAGAGTTGGACAGGA 18092
 QY 2770 CTGAGCAACTGAACCTGAACCTGACTGAAACCTTAGTAGTTTATATTACTCAGAAA 2829
 Db 18093 CTGAGTGAAGTGAACCTGAACCTGACTGAAACCTTAGTAGTTTATATTACTCAGAAA 18152
 QY 2830 ATAGTAATTCATAGTATTCATAAATTAATTCATAAATTTGGTTAAGATAATAGATTTT 2889
 Db 18153 TTGTGTTATTTGGGACATTTCTCATCTTT-ATTGAATTCCTCAAAAAAATCCAGC 18211
 QY 2890 CAAATTTGATTTTATCTTTGATTTTCTCTACTTATTTAATTTTGGGATTTTAACTATT 2949
 Db 18212 TAGAATTTGATTTGATTTTATTTATTTAATTTAATTTGATTTTAACTATT 18271
 QY 2950 CTTCAATGACTGTATTTCTAATTTTACTTATTTCTATTTTACT------TTAATGCACT 3004
 Db 18272 TTATAATACTGGGTGGTTCCCAACCAAGACCTTTCCATGTTTACAGATCATCTTCTGTGT 18331
 QY 3005 TATTTTATTTGATTTTCTAATAAATCCAGTCTCTGTTTGTTTTAAAGACATTTAAAT 3064
 Db 18332 CTTTAGTAGTGTTCAGTTCTTTTCCCACTTCTTCTTTTGTAGTTAAGTCATCC 18391
 QY 3065 TATTAATTTCTTTTAGTGTTTTACCAGTCTTTTCCAGGCTACTTCT 3110
 Db 18392 TAGATACTTTACAGTATGATGAGAAAGTTGATCTGACTCTTTGT 18437

RESULT 29

ABV99702
 ID ABV99702 standard; cDNA; 31412 BP.

XX
 AC ABV99702;

XX
 DT 12-FEB-2003 (first entry)

XX
 XX Sheep BSE-resistant prion protein cDNA.

XX
 DE Transmissible spongiform encephalopathy; neuroprotective; prion protein;
 KW bovine spongiform encephalopathy; transgenic; BSE; bovine; cervid; PrP;
 KW TSE; gene; ss; sheep.

XX
 OS Ovis aries.

XX
 PN WO200279416-A2.

XX
 PD 10-OCT-2002.

XX
 PF 28-MAR-2002; 2002WO-US0009652.

XX

PR 30-MAR-2001; 2001US-0280549P.
 (TEXA) UNIV TEXAS A & M SYSTEM.
 PA Dunne PW, Piedrahita J;
 PI WPI; 2003-092895/08.
 XX P-PSDB; ABP57901.
 XX New transgenic bovine and cervid useful for producing animals which are
 PT resistant to bovine spongiform encephalopathy and transmissible
 PT spongiform encephalopathy disease, comprise a transgene encoding a mutant
 PT PrP polypeptide.
 XX
 PS Disclosure; Page 84-93; 98pp; English.
 XX
 CC The invention relates to a novel transgenic bovine/cervid comprising a
 CC transgene encoding a mutant prion protein (PrP) polypeptide, in which a
 CC substitution has been made at position 171 of the sequence, which renders
 CC the bovine/cervid resistant to bovine spongiform encephalopathy (BSE) and
 CC transmissible spongiform encephalopathy (TSE) disease, respectively. The
 CC transgene of the invention has neuroprotective activity. The method is
 CC useful for producing a transgenic bovine or cervid resistant to BSE and
 CC TSE diseases. The bovine prion gene is useful for producing transgenic
 CC cattle exhibiting resistance to bovine spongiform encephalopathy. The
 CC sequence represents the transgene encoding the mutant PrP polypeptide in
 CC sheep
 XX
 SQ Sequence 31412 BP; 8889 A; 6517 C; 6464 G; 9542 T; 0 U; 0 Other;
 Query Match 9.6%; Score 344.199; DB 1; Length 31412;
 Best Local Similarity 69.7%; Pred. No. 6.9;
 Matches 492; Conservative 0; Mismatches 208; Indels 6; Gaps 2;
 QY 2410 TGAAGAAATGATGCTTTTGAACCTGCTGTGTGGAGAGACTCTTGAGAGTCCCTTGACT 2469
 Db 17733 TGATACAGTAGGTTTTTGAACCTGCTGTGTGGAGAGACTCTTGAGAGTCCCTTGACT 17792
 QY 2470 GCAGGAGATCCACCACTGCTCTGAAGAGATCAGCCCTGGGATTTCTTGGAGGA 2529
 Db 17793 GCAAGGAGATCCCAACCGCTCATTTCTAAGAGAGATCAGCCCTGGAATTTCTTGGAGAA 17852
 QY 2530 ATGATGCTAAAGCTGAAACTCCAGTACTTTGGCCACTGATCAGAGAGTCACTGACTG 2589
 Db 17853 ATGATGCTAAAGCTGAAATTCAGTACTTTGGCCACTGATCAGAGAGTCACTGACTG 17912
 QY 2590 GAAAGACCTGATGCTGGAGGATTTGGGGCAGGAGAGAGGGGACGACAGAGATG 2649
 Db 17913 GAAAGACTCTGATGCTGGAGGATTTGGGGCAGGAGAGAGGGGACGACAGAGATG 17972
 QY 2650 AGATGGCTGATGSCATCACTGCTGATGAGAGTCACTGCTGGTCACTCTGAGTTG 2709
 Db 17973 AGATGGCTGATGSCATCACTGCTGATGAGAGTCACTGCTGGTCACTCTGAGTTG 18032
 QY 2710 GTGATGACAGGAGGAGCTGCTGCTGGGGGATTCATGGGGTCAAAAGAGTTGGACACGA 2769
 Db 18033 TTGATGACAGGAGGAGCTGCTGCTGGGGTCAAAAGAGTTGGACACGA 18092
 QY 2770 CTGAGCACTGACTCACTGACTGACTGACTGAAACCTTAGTATTAATTAATCAAGAA 2829
 Db 18093 CTGAGTCACTGACTCACTGACTGACTGACTGAAACCTTAGTATTAATTAATCAAGAA 18152
 QY 2830 ATAGTAATTTTATGATTAATTAATTAATTTTAAATTTTAAATTAATTAATTTT 2889
 Db 18153 TTGTTATTTTGGACATTTCTTCACTTTT-ATTGAATTCCTCAAAATAATCCAGC 18211
 QY 2890 CAAATTTGATTTTATCTTTTCTACTTATTTAAATTTTGGGATTTTAACTATTT 2949
 Db 18212 TAGAATTTTGTGTCATTTATTTATTTATTAATTAATTAATTTTGGGATTTTAACTATTT 18271
 QY 2950 CTTCAATGACTTTGATTTTCTAATTTTACTTTTCTTATTTTACTTTTCTTATTTT 3004
 Db 18272 TTATAATACCTGGTTCACACCAAGACCTTTTCAATGTTTACAGATCACTTTCTGTTT 18331

QY 3005 TATTTTATGATTTTCTTAATAAAATCCAGTCTCTTGTCTTTTAAAGACTTTAAAT 3064
 Db 18332 CTTTAGTAGTGTTCAGTTCTTTTCCAGTCTTACTTTCTTTTGGTAAAGTCAATCC 18391
 QY 3065 TATTAATTTCTCTTTAGTGTCTTTTACCAGTCTCTTCCAGGCTACTTCT 3110
 Db 18392 TAGATACCTTTACAGTATGATAGTGAAGTTGTATCTGACTCTTTGT 18437
 RESULT 30
 AAD26491/c
 ID AAD26491 standard; DNA; 11034 BP.
 XX
 AC AAD26491;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Sheep alpha (1, 3) galactosyltransferase post-exon 4 DNA #2.
 XX
 KW Sheep; xenotransplantation; Gal-alpha (1,3)Gal; GAL determinant; surgery;
 KW alpha(1,3) galactosyltransferase; alpha (1, 3) GT; drug screening;
 XX gene therapy; ds.
 XX
 OS Ovis sp.
 XX
 PN WO200186096-A2.
 XX
 PD 22-NOV-2001.
 XX
 PF 14-MAY-2001; 2001WO-US015765.
 XX
 PR 15-MAY-2000; 2000US-0204148P.
 PR 13-JUN-2000; 2000US-00593316.
 XX
 XX (GERO-) GERON CORP.
 XX
 PI Denning C, Clark J;
 XX
 DR WPI; 2002-089848/12.
 XX
 PT New ovine tissue devoid of antibody-detectable Gal-alpha(1,3)Gal
 PT determinants, for xenotransplantation, and in the treatment of the human
 PT body by surgery or therapy.
 XX
 PS Claim 18; Page 76-79; 86pp; English.
 XX
 CC The patent discloses immunologically compatible animal tissue, suitable
 CC for xenotransplantation into human patients. The invention also relates
 CC to an ovine tissue devoid of antibody-detectable Gal-alpha (1,3)Gal (GAL)
 CC determinants which are made by alpha(1,3) galactosyltransferase (GT). The
 CC ovine tissue is useful for treatment of human body by surgery or therapy
 CC and in xenotransplantation, by transplanting the ovine tissue into a
 CC mammal having circulating antibody against Gal alpha(1,3)GT determinants.
 CC polynucleotide constructs of the invention are useful for inactivating an
 CC alpha(1,3) GT gene in an ovine cell. Alpha (1,3) GT sequences are useful
 CC for drug screening and for the production of GAL containing synthetic
 CC oligosaccharides. Sequences of the invention are also useful in gene
 CC therapy. The present sequence is alpha (1,3) GT post-exon 4 DNA
 XX
 SQ Sequence 11034 BP; 3010 A; 2426 C; 2435 G; 2972 T; 0 U; 191 Other;
 Query Match 9.6%; Score 342.4; DB 1; Length 11034;
 Best Local Similarity 77.2%; Pred. No. 20;
 Matches 455; Conservative 0; Mismatches 56; Indels 78; Gaps 4;
 QY 2282 CCAACCTAGATAGCATATTGAAAGCAGAGACATTACCTTGCACAAAGCCCCATCTA- 2340
 Db 4661 CTAATGTTTATAGCATATTCAAGAGAGAGACATTACTTGGCAACAAAGTCCATCTAT 4602
 QY 2341 GTCAAGCTATGTTTTTCCAGTGTCTATGATGATGAGATGAGATGGAAGAA 2400
 Db 4601 GTCAAGCTATGTTTTTCCAGTGTCTATGATGATGAGATGGAAGATGGAAGAA 4542

DT XX 20-FEB-2003 (first entry)
 DE XX Bovine EST associated with lactation/muscle/fat deposition #1308.
 XX KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
 KW muscle deposition; fat deposition; genome mapping; gene identification;
 XX KW gene analysis; cattle breeding.
 XX OS Bos Taurus.
 XX PN US2002137139-A1.
 XX PD 26-SEP-2002.
 XX PF 24-SEP-2001; 2001US-00960352.
 PR 12-JAN-1999; 99US-0115707P.
 PR 11-JAN-2000; 2000US-00480902.
 XX PA (BYAT/) BYATT J C.
 PA (MATH/) MATHIALAGAN N.
 PA (TACN/) TAO N.
 PA (WARR/) WARREN W C.
 PI Byatt JC, Mathialagan N, Tao N, Warren WC;
 DR WPI; 2003-110599/10.
 XX New nucleic acid associated with lactation, and muscle and fat
 PT deposition, useful for genome mapping, gene identification and analysis,
 PT cattle breeding, or for genetically improving cattle.
 XX Claim 2; SEQ ID NO 1308; 245bp; English.
 XX The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMFD), derived from
 CC cattle, and the LMFD nucleic acid can specifically hybridize to a second
 CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
 CC appearing as ABX34836-ABX49947, or complements of them. Also included are
 CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
 CC acid linked to a promoter and a 3' non-translated sequence that
 CC functions in the cell to cause termination of transcription and addition
 CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
 CC (2) determining a level or pattern of a molecule in a bovine cell or
 CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
 CC of the 15112 nucleic acid sequences or its complement or fragment) with a
 CC complementary nucleic acid molecule obtained from the bovine cell or
 CC tissue, where hybridization between the marker nucleic acid and the
 CC complementary nucleic acid permits the detection of the molecule; and (b)
 CC detecting the level or pattern of the complementary nucleic acid, where
 CC the detection of the complementary nucleic acid is predictive of the
 CC level or pattern of the molecule. The LMFD nucleic acid is used for
 CC determining a level or pattern of a molecule in a bovine cell or tissue.
 CC It is useful for genome mapping, gene identification and analysis, cattle
 CC breeding, preparation of constructs for use in cattle gene expression, or
 CC for genetically improving cattle. The present sequence is one of the
 CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
 CC present sequence was not shown in the specification but was obtained in
 CC electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?DocID=20020137139
 XX Sequence 406 BP; 95 A; 138 C; 77 G; 95 T; 0 U; 1 Other;
 XX
 Query Match 9.3%; Score 333; DB 1; Length 406;
 Best Local Similarity 91.9%; Pred. No. 5.6e-02;
 Matches 351; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
 QY 2405 AGCATTGAAGATGATGCTTTTGAATCTGTGGTGTGGAGAGACTCTTGGAGTCCCTT 2464
 DB 406 AGTCCCGAAGAAATGATGCTTTTGAATCTGTGGTGTGGAGAGACTCTTGGAGTCCCTT 347
 QY 2465 GGACTCGAAGAGATCCACCACTGATCTGAGAGATGAGCCCTGGGATTTCTTGG 2524

Db 346 GGACTCGAAGAGATTCACACCAAGTCCATTCGAGAGATCAGCCCTGGGATTTCTTGG 287
 QY 2525 AAGGAATGATGTAAAGCTGAAACTCCAGTACTTTGGCCACCTGATCAGAGAGCTGACT 2584
 Db 286 AGGAATGATGTGAAAGATGAAACTCCAGTACTTTGGCCACCTCATGCGAAGAGCTGGCT 227
 QY 2585 CACTGAAAAGACCCCTGATGCTGGGAGGATTTGGGGCAGGAGGAGAGAGGGGACGACAGA 2644
 Db 226 CATTGAAAAGACTCTGATGCTGGGAGGATTTGGGGCAGGAGGAGAGAGGGGACGCCAGA 167
 QY 2645 GGATGAGATGGCTGGATGGCATCACTGACTCGATGACCTGAGTCTGGTGAACCTCTGG 2704
 Db 166 GGATGAGATGGCTGGATGGCATCGCTGACTCGATGACCTGAGTCTGAGTGAACCTCCGG 107
 QY 2705 AGTTGATGATGACAGGGAGGCTCTCTCGGCGGATTCATCGGCTCAAAAGATTGGA 2764
 Db 106 AGTTGATGATGACAGGGAGGCTCTCTCGGCTGTGATTCATGGGTCGCAAGAGTCCGA 47
 QY 2765 CACGACTGAGCAACTCAACTGA 2786
 Db 46 TAGGGCTGAGCGACTGA 25
 RESULT 33
 AAH21026
 ID AAH21026 standard; DNA; 1529 BP.
 XX AC AAH21026;
 XX DT 31-AUG-2001 (first entry)
 XX DE Bovine-derived DNA fragment contig1 487_67.
 XX KW Bovine; digital DNA signature; breeding; animal product origin;
 KW identification; genetic association; breed; population; race; ds.
 OS Bos taurus.
 PN DE19959751-A1.
 PD 13-JUN-2001.
 PF 11-DEC-1999; 99DE-01059751.
 PR 11-DEC-1999; 99DE-01059751.
 PA (FRIE/) FRIES H R.
 PA (DURS/) DURSTEWITZ G.
 PI Fries HR, Durstewitz G;
 WPI; 2001-376309/40.
 PT New bovine genomic DNA sequences, useful for establishing genetic
 PT signatures, e.g. for breeding control, contain specific variable
 PT positions.
 XX Claim 1; Page 23-24; 26pp; German.
 XX This invention describes bovine DNA sequences (A) which are used in a
 CC method to establish a digital, standardized DNA signature. DNA signatures
 CC established from (A) are used to monitor breeding; to determine origin of
 CC animal products; to identify individual animals; to study genetic
 CC association and to establish signatures that are specific at the level of
 CC breed, population or race. The method is based on individual base
 CC exchanges in DNA, and these are inherited more stably (by an order of
 CC magnitude) than conventional markers. It is suitable for population-wide
 CC studies (as a high throughput test) and the presence of unequivocal
 CC sequences flanking the variable position provides a built-in
 CC standardization feature
 XX Sequence 1529 BP; 589 A; 299 C; 239 G; 358 T; 0 U; 44 Other;

```
Query Match          9.1%; Score 326.3; DB 1; Length 1529;
Best Local Similarity 73.7%; Pred. No. 1.6e+02;
Matches 452; Conservative 0; Mismatches 152; Indels 9; Gaps 3;

QY 1678 AGATTCGAGAAACATCTATTTCCTTTATGACTATGCAAAAGCCCTTCACTGTGG 1737
D 518 AGAACTACCTCTAGAAACATCTACCTGTTCAITGACAACTTAAAGACTTTGACAAATGT 577
QY 1738 GGGTCACAATAAATCTGGAAATCTTGAAGGATGGGAATACCCAGACCACTGACCTG 1797
D 578 GGATCATGCCAAGCTGGAAATCTTTAAAGAGATGGGAATACCCAGACCACTTCACTG 637
QY 1798 ACTCTGAAAAATTTGATGAGCTCAGGAAGCAACAGTT-----AGAACTGGACATGGA 1852
D 638 TCTCTCAGAAAACCTGTATGCAAGACAAAGTTGCAACAGAGTTGTATAGAACAACTGTATAGA 697
QY 1853 ACACACACTGGTTCCTCAAGTAGGAAAGAGATGTCGAAGCTGTATATTTGTCACCGGC 1912
D 698 ACAACTACTGGTTTCAAGATTTGAGAAAGGAGTCAACACAGACTGTTTACTGTCAACCACT 757
QY 1913 TTGTTTAACTTCTATGACAGACATCATGAGAAACGCTGGCTGGAAAGACACAAAGCTG 1972
D 758 TTATTTAACT--TACAGAGCACATATGTAATGCGGGCTGAATGAGTTACAGCTG 815
QY 1973 GAATCAAGATTCGGGAGAAATAGCAATTAACCTCAGATATGAGATGATACCAACCTTAA 2032
D 816 GAATCAAGATTCGGGAGAAATATCAACAAACAGATATGAGAGATACCACTTAA 875
QY 2033 TGGCAGAAAGTGAAGAGAACTAAAGACCTCTTGATGAAGGTCAAGAGAGAGAGTGAA 2092
D 876 TGGCAGAAAGTGAAGAGAACTAAAGAACCTCTTGATGAAGGTGAAGAGAGAGTGCAAG 935
QY 2093 AAGTTGGCTTAAAGCTCAACATTCAGAAACGAAAGATCATGGCATCTGTCCTCATCACTT 2152
D 936 AAACAGCTTAAACACATATTCAAAAAACTAAGATCATGGCATATGGTCCCATCACTT 995
QY 2153 CATGGGAAATAGATGGGAAACAGTGAACAGTGTGAGACTTTATTTTGGGGGCTC 2212
D 996 TACAGCAAAAAGAGGGGAAAGGTGGAAGCAATGACAAATTTCTCTTCT--TGGGCTC 1053
QY 2213 CAAATCACTGCAGATGGTGACTCAGCCATGAAATTAAGACACTTACTCTCTTGAAG 2272
D 1054 TAAATCACTGTGATGGTGACTCAGACTTGAATTTAGAAACAAATGCTTCTTTCAG 1113
QY 2273 AAAGTTAAACAA 2285
D 1114 GAACTAATAAAA 1126

RESULT 34
ABX43146/c
ID ABX43146 standard; cDNA; 423 BP.
XX AC ABX43146;
XX DT 20-FEB-2003 (first entry)
XX DE Bovine EST associated with lactation/muscle/fat deposition #8311.
XX KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
XX KW muscle deposition; fat deposition; genome mapping; gene identification;
XX KW gene analysis; cattle breeding.
XX OS Bos Taurus.
XX PN US2002137139-A1.
XX PD 26-SEP-2002.
XX PF 24-SEP-2001; 2001US-00960352.
XX PR 12-JAN-1999; 99US-0115707P.
PR 11-JAN-2000; 2000US-00480902.
```

(BYAT/) BYATT J C.
(MATH/) MATHIALAGAN N.
(TAON/) TAO N.
(WARR/) WARREN W C.

Byatt JC, Mathialagan N, Tao N, Warren WC;

WPI; 2003-110599/10.

New nucleic acid associated with lactation, and muscle and fat deposition, useful for genome mapping, gene identification and analysis, cattle breeding, or for genetically improving cattle.

Claim 2; SEQ ID NO 8311; 245pp; English.

The invention relates to a purified nucleic acid molecule associated with lactation or muscle and fat deposition (designated LMFD), derived from cattle, and the LMFD nucleic acid can specifically hybridize to a second nucleic acid molecule comprising any of 15112 nucleotide sequences, appearing as ABX4836-ABX4947, or complements of them. Also included are: (1) a transformed cell having a nucleic acid comprising an LMFD nucleic acid linked to a promoter and a 3' non-translated sequence and that functions in the cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and (2) determining a level or pattern of a molecule in a bovine cell or tissue comprising: (a) incubating a marker nucleic acid (comprising any of the 15112 nucleic acid sequences or its complement or fragment) with a complementary nucleic acid molecule obtained from the bovine cell or tissue, where hybridisation between the marker nucleic acid and the complementary nucleic acid permits the detection of the molecule; and (b) detecting the level or pattern of the complementary nucleic acid, where the detection of the complementary nucleic acid is predictive of the level or pattern of the molecule. The LMFD nucleic acid is used for determining a level or pattern of a molecule in a bovine cell or tissue. It is useful for genome mapping, gene identification and analysis, cattle breeding, preparation of constructs for use in cattle gene expression, or for genetically improving cattle. The present sequence is one of the 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The present sequence was not shown in the specification but was obtained in electronic format from the USPTO web site:
seqdata.uspto.gov/sequence.html?DocID=20020137139

Sequence 423 BP; 93 A; 128 C; 72 G; 130 T; 0 U; 0 Other;

Query Match 9.1%; Score 324; DB 1; Length 423;

Best Local Similarity 86.7%; Pred. No. 5.6e+02;

Matches 357; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 2380 GAGAGTTGCACTGTGAAGAAAGCTGAGCACTGAAGAATGTGCTTTGAAGTGGTGT 2439

D 423 GAGAGTTGCACTGTGAAGAAAGCTGAGCGCCAAAGAAATGTGCTTTGAAGTGGTGT 364

QY 2440 TGGAGAAGACTCTTGAGAGTCCCTTTGGACTGCAAGAGATCCAAACAGTCCATTCTGAAG 2499

D 363 TGGAGAAGACTCTTGAGAGTCCCTTTGGACTGCAAGAGATCCAAACAGTCCATTCTGAAG 304

QY 2500 GAGATCAGCCCTGGGATTTCTTTGGAAGGAATGATGCTAAAGCTGAACTCCAGTACTTT 2559

D 303 GAGATCAGCCCTGGGATTTCTTTGGAAGGAATGATGCTGAAAGCTGAACTCCAAACACTTT 244

QY 2560 GGCCACCTGATCAGAAAGAGCTGACTCACTGGAAAAAGACCCCTGATGCTGGAGGGGATTGGG 2619

D 243 GGTCACTGATGGGAAAACAACTGACTGTCTGAAAAGACCCCTGATGCTGGGAAAGATTGAA 184

QY 2620 GGCAGAGGAGAAAGGGGACGACAGAGGATGAGATGGCTGGATGGGATCACTGACTCGATG 2679

D 183 GGCAGAGGAGAAAGGGGACGACAGAGGATGAGATGGTGGATAGCATCACCTACTCTATG 124

QY 2680 GACGTGAGTCTGGGTGAACCTCCTGGAGTTGGTGTGAGGACAGGGAGGSCCTGTCTCTCGCGG 2739

D 123 AACTTGAGTTTGAATAAGCCCCAGGAGTTGGTGTGAGGACAGGGAAAGCATAGCATACTGCA 64

QY 2740 ATTCTGGGGTGCACAAAGAGTTGACGACGACTGACCACTGAACCTGAACCTGA 2791
 Db 63 GTCCATGGGGTGTCAAAGAGTTGACACAACTAAGTGAAGTGAACCTGAACCTGA 12

RESULT 35

ABX48669
 ID ABX48669 standard; cDNA; 410 BP.

XX AC ABX48669;

XX DT 21-FEB-2003 (first entry)

XX DE Bovine EST associated with lactation/muscle/fat deposition. #13834.

XX KW Bovine; ss; EST: expressed sequence tag; lactation; LMFD;

XX KW muscle deposition; fat deposition; genome mapping; gene identification;

XX KW gene analysis; cattle breeding.

XX OS Bos Taurus.

XX FN US2002137139-A1.

XX PD 26-SEP-2002.

XX PF 24-SEP-2001; 2001US-00960352.

XX PR 12-JAN-1999; 99US-0115707P.

XX PR 11-JAN-2000; 2000US-00480902.

XX PA (BYAT/) BYATT J C.

XX PA (MATH/) MATHIALAGAN N.

XX PA (TAON/) TAO N.

XX PA (WARR/) WARREN W C.

XX PI Byatt JC, Mathialagan N, Tao N, Warren WC;

XX DR WPI; 2003-110599/10.

XX PT New nucleic acid associated with lactation, and muscle and fat

XX PT deposition, useful for genome mapping, gene identification and analysis,

XX PT cattle breeding, or for genetically improving cattle.

XX PS Claim 2; SEQ ID NO 13834; 245pp; English.

XX CC The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMFD), derived from
 CC cattle, and the LMFD nucleic acid can specifically hybridise to a second
 CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
 CC appearing as ABX34836-ABX49947, or complements of them. Also included are
 CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
 CC acid linked to a promoter and a 3' non-translated sequence that
 CC functions in the cell to cause termination of transcription and addition
 CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
 CC (2) determining a level or pattern of a molecule in a bovine cell or
 CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
 CC of the 15112 nucleic acid sequences or its complement or fragment) with a
 CC complementary nucleic acid molecule obtained from the bovine cell or
 CC tissue, where hybridisation between the marker nucleic acid and the
 CC complementary nucleic acid permits the detection of the molecule; and (b)
 CC detecting the level or pattern of the complementary nucleic acid, where
 CC the detection of the complementary nucleic acid is predictive of the
 CC level or pattern of the molecule. The LMFD nucleic acid is used for
 CC determining a level or pattern of a molecule in a bovine cell or tissue.
 CC It is useful for genome mapping, gene identification and analysis, cattle
 CC breeding, preparation of constructs for use in cattle gene expression, or
 CC for genetically improving cattle. The present sequence is one of the
 CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
 CC present sequence was not shown in the specification but was obtained in
 CC electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?docID=20020137139

XX SQ Sequence 410 BP; 132 A; 67 C; 104 G; 106 T; 0 U; 1 Other;

Query Match 9.0%; Score 321.6; DB 1; Length 410;
 Best Local Similarity 88.9%; Pred. No. 5.9e+02;
 Matches 368; Conservative 0; Mismatches 40; Indels 6; Gaps 2;

QY 2242 ATGAAATTAAGACACACTTACTCTCTTGGAGAAAGTTA--ACCACCTAGATAGCATAT 2299
 Db 1 ATGAAATTAAGACATGCTTACTCTCTTGGAGAAAGTTATGACCAACCTAGATAGCATAT 60
 QY 2300 TGAAGAAGCAGACACATTACCTTTGCCAACAAAGCCCCCATCTAGTCAAGGCTATGGTTTTTC 2359
 Db 61 TCAAGAAGCAGACACATTACTTTTGGCCAAACAAGGTCCTGTAGTCAAGGCTATGGTTTTTC 120
 QY 2360 CAGTGGTCATGTATGGATGTGAGAGTTGGAGCTGTGAAGAAAGCTGAGCACTGAAGAATTG 2419
 Db 121 CAGTAGTCATGTACGATGTGAGAGTTGGAGCTGTGAAGAAAGCTGAGCACTGAAGAATTG 180
 QY 2420 ATGCTTTTGAAGCTGTGGTGGAGAGACTCTTGAGAGTCCCTTGGAGCTGCAAGGAGAT 2479
 Db 181 ATGCTTTTGAAGCTGTGGTGGAGAGACTCTTGAGAGTCCCTTGGAGCTGCAAGGAGAT 240
 QY 2480 CCAACGAGTCCATTTCTGAAGGAGATCAGCCCTGGGATTTCTTTGGAAGGAATGATCTAA 2539
 Db 241 CCAACGAGTCTATTCTTAAAGGAGATCAGTCTCTG---TTCITTTGGAAGGACTGAAGCTAA 296
 QY 2540 AGCTGAAACTCCAGTACTTTTGGCCACCTGATCAGAGAGCTGACTCCTGGAAGACCC 2599
 Db 297 AGCTGAAACTCAAATTTCTTTGGCCACCTGATCAGAGAGCTGACTCCTTGAAGAGACTT 356
 QY 2600 TGATGCTGGAGGGATTGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGATGAT 2653
 Db 357 TGATGCTTGGAAAGACTGAAGGCTGGAAGAGAGAGGAGGATGACAGANGATGAT 410

RESULT 36

ABX40334/c

ID ABX40334 standard; cDNA; 423 BP.

XX AC ABX40334;

XX DT 20-FEB-2003 (first entry)

XX DE Bovine EST associated with lactation/muscle/fat deposition #5499.

XX KW Bovine; ss; EST: expressed sequence tag; lactation; LMFD;

XX KW muscle deposition; fat deposition; genome mapping; gene identification;

XX KW gene analysis; cattle breeding.

XX OS Bos Taurus.

XX PN US2002137139-A1.

XX PD 26-SEP-2002.

XX PF 24-SEP-2001; 2001US-00960352.

XX PR 12-JAN-1999; 99US-0115707P.

XX PR 11-JAN-2000; 2000US-00480902.

XX PA (BYAT/) BYATT J C.

XX PA (MATH/) MATHIALAGAN N.

XX PA (TAON/) TAO N.

XX PA (WARR/) WARREN W C.

XX PI Byatt JC, Mathialagan N, Tao N, Warren WC;

XX DR WPI; 2003-110599/10.

XX PT New nucleic acid associated with lactation, and muscle and fat

XX PT deposition, useful for genome mapping, gene identification and analysis,

XX PT cattle breeding, or for genetically improving cattle.

XX PS Claim 2; SEQ ID NO 5499; 245pp; English.

XX The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMFD), derived from
CC cattle, and the LMFD nucleic acid can specifically hybridize to a second
CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
CC appearing as ABX34836-ABX49947, or complements of them. Also included are
CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
CC acid linked to a promoter and a 3' non- translated sequence that
CC functions in the cell to cause termination of transcription and addition
CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
CC (2) determining a level or pattern of a molecule in a bovine cell or
CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
CC of the 15112 nucleic acid sequences or its complement or fragment) with a
CC complementary nucleic acid molecule obtained from the bovine cell or
CC tissue, where hybridisation between the marker nucleic acid and the
CC complementary nucleic acid permits the detection of the molecule; and (b)
CC detecting the level or pattern of the complementary nucleic acid, where
CC the detection of the complementary nucleic acid is predictive of the
CC level or pattern of the molecule. The LMFD nucleic acid is used for
CC determining a level or pattern of a molecule in a bovine cell or tissue.
CC It is useful for genome mapping, gene identification and analysis, cattle
CC breeding, preparation of constructs for use in cattle gene expression, or
CC for genetically improving cattle. The present sequence is one of the
CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
CC present sequence was not shown in the specification but was obtained in
CC electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?DocID=20020137139
XX
SQ Sequence 423 BP; 93 A; 128 C; 72 G; 130 T; 0 U; 0 Other;

Query Match 9.0%; Score 320.8; DB 1; Length 423;
Best Local Similarity 86.2%; Pred. No. 5.7e+02;
Matches 355; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

Qy 2380 GAGAGTTGGAGTGTGAAGAAGCTGACGACCTGAAGAAATGATGCTTTTGAACCTGTGGTGT 2439
Db 423 GAGAGTTGGAGTGTGAAGAAGCTGAGCGCCAAAGAAATGATGCTTCTGAACCTGTGGTGT 364
Qy 2440 TGGAGAGACTCTGAGAGTCCCTTGGAGTGCAGAGAGATCCACAGTCCATCTCTGAAG 2499
Db 363 TGGAGAGACTCTTGAAGATCCCTTGGAGTGCAGAGAGATCCACAGTCCATCTCTGAAG 304
Qy 2500 GAGATCAGCCCTGGGATTTCTTGGAAAGGAATGATGCTAAAGCTGMAAATCCAGTACTTT 2559
Db 303 GAGATCAGCCCTGGGATTTCTTGGAAAGGAATGATGCTAAAGCTGMAAATCCACACTTT 244
Qy 2560 GGCACCTGATCAGAGAGCTGACTCTGTAAGAAGCCCTGATGCTGGAGGAGATTGG 2619
Db 243 GGTCACTGATGGGAAAACCTGACTGATCTGAAAAGACCCTGATGCTGGGAAAGATTGAA 184
Qy 2620 GGCAGGAGGAAGGGACGACAGAGGATGAGATGGCTGGATGGCATCACTGACTCGATG 2679
Db 183 GCGAGGAGGAGAGGGGACGACAGAGGATGAGATGGCTGGATGGCATCACTGACTCTATG 124
Qy 2680 GACGTGAGTCTGGGTGAACCTCTGAGTTGGTGTGATGACAGGGAGGCTCTCTCTGGGGG 2739
Db 123 AACTTGATTTGAATTAAGCCCGAGGTGGTGTGATGACAGGGAGGATGACATGCTACTGCA 64
Qy 2740 ATTATGGGTCACAAAGAGTTGACAGCTGACGACTGCAACTGAAGTGAAGTGA 2791
Db 63 GTCCATGGGTTGCAAGAGTTGGACACAACTAAGTGAAGTGAAGTGAAGTGA 12

RESULT 37
ABX49143/C
ID ABX49143 standard; cDNA; 409 BP.
XX
AC ABX49143;
XX
XX
DT 21-FEB-2003 (first entry)
XX
DE Bovine EST associated with lactation/muscle/fat deposition #14308.

KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
KW muscle deposition; fat deposition; genome mapping; gene identification;
KW gene analysis; cattle breeding.
OS Bos Taurus.
XX US2002137139-A1.
PN 26-SEP-2002.
XX 24-SEP-2001; 2001US-00960352.
XX 12-JAN-1999; 99US-0115707P.
PR 11-JAN-2000; 2000US-00480902.
XX (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARR/) WARREN W C.
XX Byatt JC, Mathialagan N, Tao N, Warren WC;
PI WPI; 2003-110599/10.
DR New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and analysis,
PT cattle breeding, or for genetically improving cattle.
XX
XX Claim 2; SEQ ID NO 14308; 245pp; English.
XX The invention relates to a purified nucleic acid molecule associated with:
XX lactation or muscle and fat deposition (designated LMFD), derived from
CC cattle, and the LMFD nucleic acid can specifically hybridize to a second
CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
CC appearing as ABX34836-ABX49947, or complements of them. Also included are
CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
CC acid linked to a promoter and a 3' non- translated sequence that
CC functions in the cell to cause termination of transcription and addition
CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
CC (2) determining a level or pattern of a molecule in a bovine cell or
CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
CC of the 15112 nucleic acid sequences or its complement or fragment) with a
CC complementary nucleic acid molecule obtained from the bovine cell or
CC tissue, where hybridisation between the marker nucleic acid and the
CC complementary nucleic acid permits the detection of the molecule; and (b)
CC detecting the level or pattern of the complementary nucleic acid, where
CC the detection of the complementary nucleic acid is predictive of the
CC level or pattern of the molecule. The LMFD nucleic acid is used for
CC determining a level or pattern of a molecule in a bovine cell or tissue.
CC It is useful for genome mapping, gene identification and analysis, cattle
CC breeding, preparation of constructs for use in cattle gene expression, or
CC for genetically improving cattle. The present sequence is one of the
CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
CC present sequence was not shown in the specification but was obtained in
CC electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?DocID=20020137139
XX
SQ Sequence 409 BP; 90 A; 97 C; 84 G; 137 T; 0 U; 1 Other;

Query Match 9.0%; Score 320.4; DB 1; Length 409;
Best Local Similarity 89.3%; Pred. No. 6e+02; 42; Indels 2; Gaps 2;
Matches 366; Conservative 0; Mismatches 42; Indels 2; Gaps 2;

Qy 1748 AAACGTGGAAATTTCTGAAGGATGGGAATACACAGACCCACCTGACCTCTTGAAA 1807
Db 409 AAACGTGGCAATTTCTGAAGAGATGGGAATACACAGACCCACCTGATCTCTTGAGA 350
Qy 1808 AATTTGTATGCGAGTCAGGACGACAGTGTAGAACTGGACATGGACACAGCTGGTTC 1867
Db 349 AATCTGTATGCGAGTCAGGACGACAGTGTAGAACTGGACATGGACACAGCTGGTTC 290
Qy 1868 CAAGTAGGAAAGAGGATGTGCAAGCTGTATATGTTCACCCGGCTTGTAACTTCTAT 1927

Db	289	CAAAATAGGACAAGAAATCCGTCGAAGGTAGTATATGTGTCACCGCTGCTTATTTAACTTATAT	230
Qy	1928	GCAGAG-ACATCATGAGAAACGCTGGCTGGAAGAACACCAAGTGGAAATCAAGATTGCC	1986
Db	229	GCAGAGTCCATGATGAGAAACTCTGGCTGGAAGAAACACAAGCTGGAATCAAGATTGCC	170
Qy	1987	GGGAGAAATAGCAATAACCTCAGATATGCAGATGATACCAACCTTATGGCAGAAAGTGAA	2046
Db	169	GGGAGAAATATCAATAACCTCAGATATGCAGATGATACCAACCTTATGGCAGAAAGTGAA	110
Qy	2047	GAGGAACATAAAGAGCCCTCTTTGATCAAGGTGAAAGAGAGAGAGTGAAGAAAGTTGGCTTAAAG	2106
Db	109	GAGGAACATAAAGAGCCCTCTTTGATCAAGGTGAAAGAGAGAGTGAAGAAAGTTGGCTTAAAG	50
Qy	2107	CTCAACATTGAGAAACGAAAGATCATGGCATCTGGTCCCATCACTTCATG	2156
Db	49	CTTCACCTTCAGAAAATG-NGATCATGGCATCGGTCCTCCGTTCCATG	1
RESULT 38			
AAK56452/c			
ID AAK56452 standard; cDNA; 529 BP.			
XX	AC	AAK56452;	
XX	AC		
DT	06-NOV-2001	(first entry)	
DT			
XX	Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:1512.		
DE	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;		
XX	cytostatic; gene therapy; vaccine; metastasis; ss.		
KW			
KX			
OS	Homo sapiens.		
OS			
XX	WO200157182-A2.		
PN			
XX	09-AUG-2001.		
XX	17-JAN-2001; 2001WO-US0001354.		
XX			
PR	31-JAN-2000; 2000US-0179065P.		
PR	04-FEB-2000; 2000US-0180628P.		
PR	24-FEB-2000; 2000US-0184664P.		
PR	02-MAR-2000; 2000US-0186350P.		
PR	16-MAR-2000; 2000US-0189874P.		
PR	17-MAR-2000; 2000US-0190076P.		
PR	18-APR-2000; 2000US-0198123P.		
PR	13-MAY-2000; 2000US-0205515P.		
PR	07-JUN-2000; 2000US-0209467P.		
PR	28-JUN-2000; 2000US-0214886P.		
PR	30-JUN-2000; 2000US-0215135P.		
PR	07-JUL-2000; 2000US-0216647P.		
PR	07-JUL-2000; 2000US-0216880P.		
PR	11-JUL-2000; 2000US-0217487P.		
PR	11-JUL-2000; 2000US-0217456P.		
PR	14-JUL-2000; 2000US-0218290P.		
PR	26-JUL-2000; 2000US-0220963P.		
PR	26-JUL-2000; 2000US-0220964P.		
PR	14-AUG-2000; 2000US-0224518P.		
PR	14-AUG-2000; 2000US-0224519P.		
PR	14-AUG-2000; 2000US-0225447P.		
PR	14-AUG-2000; 2000US-0225213P.		
PR	14-AUG-2000; 2000US-0225514P.		
PR	14-AUG-2000; 2000US-0225566P.		
PR	14-AUG-2000; 2000US-0225267P.		
PR	14-AUG-2000; 2000US-0225268P.		
PR	14-AUG-2000; 2000US-0225270P.		
PR	14-AUG-2000; 2000US-0225447P.		
PR	14-AUG-2000; 2000US-0225757P.		
PR	14-AUG-2000; 2000US-0225758P.		
PR	14-AUG-2000; 2000US-0225759P.		
PR	18-AUG-2000; 2000US-0226279P.		
PR	22-AUG-2000; 2000US-0226681P.		
PR	22-AUG-2000; 2000US-0226686P.		
PR	17-NOV-2000; 2000US-0249208P.		
PR	22-AUG-2000; 2000US-0227182P.		
PR	23-AUG-2000; 2000US-0227009P.		
PR	30-AUG-2000; 2000US-0228924P.		
PR	01-SEP-2000; 2000US-0229287P.		
PR	01-SEP-2000; 2000US-0229343P.		
PR	01-SEP-2000; 2000US-0229344P.		
PR	01-SEP-2000; 2000US-0229345P.		
PR	05-SEP-2000; 2000US-0229509P.		
PR	05-SEP-2000; 2000US-0229513P.		
PR	06-SEP-2000; 2000US-0230437P.		
PR	06-SEP-2000; 2000US-0230438P.		
PR	08-SEP-2000; 2000US-0231242P.		
PR	08-SEP-2000; 2000US-0231243P.		
PR	08-SEP-2000; 2000US-0231244P.		
PR	08-SEP-2000; 2000US-0231413P.		
PR	08-SEP-2000; 2000US-0231414P.		
PR	08-SEP-2000; 2000US-0232080P.		
PR	08-SEP-2000; 2000US-0232081P.		
PR	12-SEP-2000; 2000US-0231968P.		
PR	14-SEP-2000; 2000US-0233397P.		
PR	14-SEP-2000; 2000US-0233398P.		
PR	14-SEP-2000; 2		

Db	353	GACTGGAATGGGTGAATTTAACTCAGATAAACCAATTATATCTACTACCGTGGGAGGAATC	234
Qy	723	CCTCAGAAAGAAATGGGTAGTCCCATCATGCTCAACAAAGAGTCCGAAATCAGTACTTGG	782
Db	293	CCTTAGAAAGAAACAGAGTAGCCATCATGTGTCAACAAAGAGTCTGAAATCGAGTACTTGG	234
Qy	783	ATGCAGTCTCAAAAACGACAGAGATGATCTCTGTGTGTTTCCAAGGCAAAACCATTCAAATAT	842
Db	233	ATGCAATCTCAAAAACGAAAGATGA-----CTGGTTTCCAAGGCAAAACCATTCAAATAT	180
Qy	843	CACAGTAATCCAAAGTCTATGCCCCAACCCAGTATATGCTGAAGAGCTGAAGTTGAACGGTC	902
Db	179	CACAGTAATCCAGCCCTATGCCCCAACCCAGTAAAGCTGTGAAGAGCTGAAGTTGAATGGTT	120
Qy	903	CTATGAAGACCTACAAGACCTTTTAGAACTTAACACCCCAAAAA-AGATGTGCTCTTCATTA	961
Db	119	CTATGAAGACCTACAAGACCTTTTAGAACTTAACACCCCAAAAAACAGATGTCTCTTTTCGTTA	60
Qy	962	TAGGGAGCTGGAATGCAAAAGTAGGAAGCAAAACACCTCGAGTAACAGGCAAAATTT	1020
Db	59	TGGGGAGCTGGAATGCAAAAGTAGGAAGCAAAACATCTGGGGTAAACAGAGAAATTT	1
RESULT 39			
ABX44380			
ID ABX44380 standard; cDNA; 425 BP.			
XX	AC	ABX44380;	
XX	DT	21-FEB-2003 (first entry)	
XX	DE	Bovine EST associated with lactation/muscle/fat deposition #9545.	
XX	KW	Bovine; ss; EST; expressed sequence tag; lactation; LMFD;	
XX	KW	muscle deposition; fat deposition; genome mapping; gene identification;	
XX	KW	gene analysis; cattle breeding.	
XX	OS	Bos Taurus.	
XX	PN	US2002137139-A1.	
XX	PD	26-SEP-2002.	
XX	PF	24-SEP-2001; 2001US-00960352.	
XX	PR	12-JAN-1999; 99US-0115707P.	
XX	PR	11-JAN-2000; 2000US-00480902.	
XX	PA	(BYAT/) BYATT J C.	
XX	PA	(MATH/) MATHIALAGAN N.	
XX	PA	(TAON/) TAO N.	
XX	PA	(WARR/) WARREN W C.	
XX	PI	Byatt JC, Mathialagan N, Tao N, Warren WC;	
XX	DR	WPI; 2003-110599/10.	
XX	PT	New nucleic acid associated with lactation, and muscle and fat	
XX	PT	deposition, useful for genome mapping, gene identification and analysis,	
XX	PT	cattle breeding, or for genetically improving cattle.	
XX	PS	Claim 2; SEQ ID NO 9545; 245bp; English.	
XX	CC	The invention relates to a purified nucleic acid molecule associated with	
XX	CC	lactation or muscle and fat deposition (designated LMFD), derived from	
XX	CC	cattle, and the LMFD nucleic acid can specifically hybridize to a second	
XX	CC	nucleic acid molecule comprising any of 1512 nucleotide sequences,	
XX	CC	appearing as ABX34836-ABX49947, or complements of them. Also included are	
XX	CC	(1) a transformed cell having a nucleic acid comprising an LMFD nucleic	
XX	CC	acid linked to a promoter and a 3' non-translated sequence that	
XX	CC	functions in the cell to cause termination of transcription and addition	
XX	CC	of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and	
XX	CC	(2) determining a level or pattern of a molecule in a bovine cell or	

CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
 CC of the 15112 nucleic acid sequences or its complement or fragment) with a
 CC complementary nucleic acid molecule obtained from the bovine cell or
 CC tissue, where hybridisation between the marker nucleic acid and the
 CC complementary nucleic acid permits the detection of the molecule; and (b)
 CC detecting the level or pattern of the complementary nucleic acid, where
 CC the detection of the complementary nucleic acid is predictive of the
 CC level or pattern of the molecule. The LMFD nucleic acid is used for
 CC determining a level or pattern of a molecule in a bovine cell or tissue.
 CC It is useful for genome mapping, gene identification and analysis, cattle
 CC breeding, preparation of constructs for use in cattle gene expression, or
 CC for genetically improving cattle. The present sequence is one of the
 CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
 CC present sequence was not shown in the specification but was obtained in
 CC electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?DocID=20020137139
 XX
 SQ Sequence 425 BP; 144 A; 70 C; 100 G; 111 T; 0 U; 0 Other;

Query Match 8.8%; Score 313.4; DB 1; Length 425;
 Best Local Similarity 87.4%; Pred. No. 6e+02;
 Matches 375; Conservative 0; Mismatches 46; Indels 8; Gaps 3;
 QY 2073 GGTGAAGAGAGAGTGAAGAAGTTGGCTTAAGCTCAACATTCAGAAACGAGATCAT 2132
 Db 3 GGTGAAGAGAGAGTGAAGAAGTTGGCTTAATCTCAACATTCAGAAACGAGATCAT 62
 QY 2133 GGCATCTGGTCCCATCACTTCATCGGAAATAGATGGGAAACAGTGGAAACAGTGTGAGA 2192
 Db 63 GGCAGCTGATCCCATTAATTTTCATGGCAATAGATGGGATACAAATGGAAACAGTGAGAGA 122
 QY 2193 CTTTATTTTGGGGGCTCCAAATCACTGCAGATGGTCACTGCAGCCATGAATTAATAA 2252
 Db 123 CTTTATGTTTT--TGGGCTCCAAATCACTGCAGATGGTGAATGCAGCCATGAATTAATAA 180
 QY 2253 AGACACTTACTCTTGGAGAAAGTTA--ACCAACCTAGATAGCATATTTGAAGAGCAGA 2310
 Db 181 TGACTCTTCTTGGAGAAAGCTATGACCGAGCTAGACAGCATATTTAAAGAGCAGA 240
 QY 2311 GACATTAACCTTGCACAAAGCCCATCTAGTCAAGCTATGGTTTTTCCAGTGGTCAATG 2370
 Db 241 AACATTAACCTTACCAAAAGGTCATCTATTTCAAAGCTA---TTTCTAGTAGTCAATG 296
 QY 2371 TATGGATGTGAGAGTTGGACTGTGAAGAAAGCTGAGCACTGAAGAAATTTGATGCTTTTGA 2430
 Db 297 TATGGTTGTGAGAGTTGGACTATTAAGAAAGCTGAGTCTGAAGAACTGATGTTTGA 356
 QY 2431 CTGTGGTGTGGAGAAAGCTTTGAGAGTCCCTTGGACTCCAGAGGAGATCCAAACCAAGTCC 2490
 Db 357 CTGGGTGTGGAGAAAGCTCTGAGAGTCCCTTGAATCTCAAGAGGAGATCCAAACCAAGTCA 416
 QY 2491 ATTCCTGAAG 2499
 Db 417 ATCCTAAG 425

RESULT 40
 AAH21024/c
 ID AAH21024 standard; DNA; 1435 BP.
 XX
 AC AAH21024;
 XX
 DT 31-AUG-2001 (first entry)
 XX
 DE Bovine-derived DNA fragment contig1 422_19.
 XX
 KW Bovine; digital DNA signature; breeding; animal product origin;
 XX identification; genetic association; breed; population; race; ds.
 XX
 OS Bos taurus.
 XX
 PN DE19959751-A1.
 XX

PD 13-JUN-2001.
 XX
 PF 11-DEC-1999; 99DE-01059751.
 XX
 PR 11-DEC-1999; 99DE-01059751.
 XX
 PA (FRIE/) FRIES H R.
 PA (DURS/) DURSTEWITZ G.
 XX
 PI Fries HR, Durstewitz G;
 XX WPI; 2001-376309/40.
 DR
 XX New bovine genomic DNA sequences, useful for establishing genetic
 PT signatures, e.g. for breeding control, contain specific variable
 PT positions.
 XX
 XS Claim 1; Page 22; 26pp; German.
 XX

CC This invention describes bovine DNA sequences (A) which are used in a
 CC method to establish a digital, standardized DNA signature. DNA signatures
 CC established from (A) are used to monitor breeding, to determine origin of
 CC animal products; to identify individual animals; to study genetic
 CC association and to establish signatures that are specific at the level of
 CC breed, population or race. The method is based on individual base
 CC exchanges in DNA, and these are inherited more stably (by an order of
 CC magnitude) than conventional markers. It is suitable for population-wide
 CC studies (as a high throughput test) and the presence of unequivocal
 CC sequences flanking the variable position provides a built-in
 CC standardization feature
 XX

SQ Sequence 1435 BP; 309 A; 301 C; 241 G; 584 T; 0 U; 0 Other;
 Query Match 8.7%; Score 312.4; DB 1; Length 1435;
 Best Local Similarity 74.5%; Pred. No. 1.8e+02;
 Matches 427; Conservative 0; Mismatches 116; Indels 30; Gaps 3;
 QY 1069 CHAGAAATGACCTGCTCATAGCAAAACCCCTTCCAAACACAGAGAAAGCTCTAC 1128
 Db 1079 CAAGAGACATACCTGATCAGAGCAAAACCCCTTCCAAATACACAGAGATGACTCTAC 1020
 QY 1129 ACATGGACATCACAGATGGTCAACACCGAATCAGATTGATTATTTCTTTGAGCCAA 1188
 Db 1019 ACATGGACATCACACCTGGTCAATACCGCAATCAGACTGATTATTTCTTTAGGGCTGA 960
 QY 1189 AGATGGAGAGCTCTATACAGTCAAGCAAAACACAGACGAGGCTTACTGTGGCTCAGAT 1248
 Db 959 AGATGGACAAACTCTATACAGTCAAGCAAAACACAGACCTAGAG-----TTGACTCAGAT 906
 QY 1249 CATGAACCTCTTATGGCCAAATTCAGACTTAAATTTGAAGAAAGTA--GGGAAACCTCTAG 1307
 Db 905 TATTAGCACCTCTCTGCAAAAGTCAGGCTTAAATTTGAAGAAAGTAGGGGAAACTACTAG 846
 QY 1308 ATCACTCAGGTAAGACCTTAAATCCAAATCCCTTTATGATTATATACAGTGGAAAGTAATAG 1367
 Db 845 GCCATTGAGGATGACCTAATCAATCTCTTATGATTACACAGTGGGGTGTGAACAG 786
 QY 1368 ATTTAAGGCTTAGATCTGATAGACAGAGTACCTTAATGAACCTATGGACAGAGGTTCTATG- 1426
 Db 785 ATTCAGGGATTAGATCTGGTAGACAGTATGCTCTGAAGAACTATGGAAGGAGGTTCTATGT 726
 QY 1427 -----ACATTGTACAGGACAGGATCGAGATCGAGATCCATCCCATG 1464
 Db 725 TTGTCTGTACAATAACATTGTACATTGTACAGAGGCGAGTGATCAAAACCATTCCTCAAG 666
 QY 1465 GAAAGAAATGCAAAAGCAAAATGGCTGTCTGGGAGGCGCTTACAAATAGCTGTGAAA 1524
 Db 665 AAAAGAAATGCAAAAGGCAAAAGTGTATCTACAGAGGCTTTACAAATAGCTGTGAAA 606
 QY 1525 AGAAGAGAGTGAAGAAAGCAAAAGGAAAGAAAGTAAAGCATCTGAAATGAGAGTTC 1584
 Db 605 AGAAGAGAGTGAAGCAAAAGGAAAGGAAAGGAAAGTATATACCTGACTGATATCTGTAGA 546

QY	1585	CAGAGAACTTCAGTGTTCAGCTGGTTTGTAG	1617
DB	545	CAGAGTGCCTGATGAACACTACGGATTGAGGTAG	513
RESULT 41			
ID	ABX38226	standard; cDNA; 422 BP.	
XX	XX	ABX38226;	
XX	XX	20-FEB-2003 (first entry)	
XX	XX	Bovine EST associated with lactation/muscle/fat deposition #3391.	
DE	XX	Bovine; ss; EST; expressed sequence tag; lactation; LMPD;	
KW	KW	muscle deposition; fat deposition; genome mapping; gene identification;	
KW	KW	gene analysis; cattle breeding.	
XX	XX	Bos Taurus.	
XX	XX	US2002137139-A1.	
PN	XX	26-SEP-2002.	
XX	XX	24-SEP-2001; 2001US-00960352.	
XX	XX	12-JAN-1999; 99US-0115707P.	
PR	PR	11-JAN-2000; 2000US-00480902.	
XX	XX	(BYAT/) BYATT J C.	
PA	PA	(NATH/) MATHIALAGAN N.	
PA	PA	(TAON/) TAO N.	
PA	PA	(WARR/) WARREN W C.	
XX	XX	Byatt JC, Mathialagan N, Tao N, Warren WC;	
XX	XX	WPI; 2003-110599/10.	
XX	XX	New nucleic acid associated with lactation, and muscle and fat	
XX	XX	deposition, useful for genome mapping, gene identification and analysis,	
PT	PT	cattle breeding, or for genetically improving cattle.	
XX	XX	Claim 2; SEQ ID NO 3391; 245pp; English.	

Query Match	8.6%;	Score 307.8;	DB 1;	Length 422;
Best Local Similarity	86.0%;	Pred. No. 6.2e+02;		
Matches 363;	Conservative 0;	Mismatches 57;	Indels 2;	Gaps 2;
QY	1714	CTATGCAAAAAGCCCTTTGACTGTGGGGGTACATAAACTCTGAAAAATTCGAAAGGGAT	1773	
DB	1	CTGGCTGAAAACCTTTGACTGTGTAGATCACAACATACTGTGAAAAATTGTTAAAGAGAT	60	
QY	1774	GGGAATACCGAGACCACCTGACCTGACTCTTTGAAAAATTTCTATGTCAGGTGAGGAAGCAAC	1833	
DB	61	GGGAATACCGAGAACCTTATCTCTCTCTGAGAACCTGTGTGTAGGTCAAGAGCAAC	120	
QY	1834	AGTTAGAACTGGACATGGACACACAGACTGGTTTCCAAAGTAGGAAAAAGGAGTATGTCAAGG	1893	
DB	121	AGTTAGAACCAAGGATGGAAACAACAGACTGGTTCCAACTGGGAAGGGAGCTACATCAAGG	180	
QY	1894	CTGTATATTGTACCCGGCTTGTAACTTCTATGACAG-ACATCATGAGAACCGCTGG	1952	
DB	181	CAATATATTGTCAACCCTGCTTATTTAACTTATATGCAAGGTACATCATGAGAAATGCTCG	240	
QY	1953	GCTGGAAGAAGCACAAAGCTGGAATCAAGATTGC-CGGGAGAAATAGCAATAAAGCTCAGAT	2011	
DB	241	GCTGGATGAAAACAAGCTGGAATCAAGATTGCTGGGGAGAAATATCAATAAAGCTCAGAT	300	
QY	2012	ATGCAGATCATACCAACCCCTTATGGCAGAAAGTGAAGAGGAACCTAAAAAGCCTCTTGTATGA	2071	
DB	301	ATGCAGATGGCACCCTCCTTATGGCAGAAAGCAAAAGAGGAGCTAAAAAGCCTCTTGTATGA	360	
QY	2072	AGGTGAAAAGAGAGTGATAAAAAGTTGGCTTAAAGCTCAACATTCAGAAACAGAGATCA	2131	
DB	361	AAGTGAAGAAGTGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCATTAACCTAAGATCA	420	
QY	2132	TG 2133		
DB	421	TG 422		

The invention relates to a purified nucleic acid molecule associated with lactation or muscle and fat deposition (designated LMPD), derived from cattle, and the LMPD nucleic acid can specifically hybridise to a second nucleic acid molecule comprising any of 15112 nucleotide sequences, appearing as ABX34836-ABX49947, or complements of them. Also included are acid linked to a promoter and a 3' non-translated sequence that functions in the cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and (2) determining a level or pattern of a molecule in a bovine cell or tissue comprising: (a) incubating a marker nucleic acid (comprising any of the 15112 nucleic acid sequences or its complement or fragment) with a complementary nucleic acid molecule obtained from the bovine cell or tissue, where hybridisation between the marker nucleic acid and the complementary nucleic acid permits the detection of the molecule; and (b) detecting the level or pattern of the complementary nucleic acid, where the detection of the complementary nucleic acid is predictive of the level or pattern of the molecule. The LMPD nucleic acid is used for determining a level or pattern of a molecule in a bovine cell or tissue. It is useful for genome mapping, gene identification and analysis, cattle breeding, preparation of constructs for use in cattle gene expression, or for genetically improving cattle. The present sequence is one of the 15112 bovine LMPD EST (expressed sequence tag) nucleic acids. Note: The present sequence was not shown in the specification but was obtained in electronic format from the USPTO web site: seqdata.uspto.gov/sequence.html?DocID=20020137139

Sequence 422 BP; 152 A; 73 C; 100 G; 97 T; 0 U; 0 Other;

Query Match			
Best Local Similarity 86.0%; Score 307.8; DB 1; Length 422;			
Matches 363; Conservative 0; Mismatches 57; Indels 2; Gaps 2;			
QY	1714	CTATGCAAAAGCCCTTACTGTGGGGTCAATAAACTGTGAAAAATCTGAAGGAT	1773
DB	1	CTGGCTGAAAACCTTTGACTGTGTAGATCAACAATCTGTGAAAAATGTTAAAGAGAT	60
QY	1774	GGGAATACCAAGCACCCTGACCTGACTCTTGAATAATTTGATGAGTCAGGTCAGGAAGCAAC	1833
DB	61	GGGAATACCAAGAACCTTATCTGCTTCTTGAGAAACCTGTGTGTAGGTCAAGAGCAAC	120
QY	1834	AGTTAGAACTGGACATGGAACAACAGACTGGTTCCAAAGTAGAAAAAGGATATGTCAAGG	1893
DB	121	AGTTAGAACCAAGGATGGAACAACAGACTGGTTCCAAAGTAGAAAAAGGATATGTCAAGG	180
QY	1894	CTGTATATTGTCAACCGGCTTGTAACTTCTATGTCAGAG-ACATCATGAGAAAACGGCTGG	1952
DB	181	CAATATATTGTACCCCTCTTATTTAACTTATATGAGAGTACATCATGAGAAATGCTGG	240
QY	1953	GCTGGAAGAAGCACAAAGCTGGAATCAAGATTGC-CGGGAGAAATAGCAATTAACCTCAGAT	2011
DB	241	GCTGGATGAACACAAAGCTGGAATCAAGATTGCTGGGAGAAATATCAATAACCTCAGAT	300
QY	2012	ATGCAGATGATACACCTTATGGCAGAAAGTGAAGAGAACTAAAAGCCTCTTGATGA	2071
DB	301	ATGCAGATGGCACCCTCTTATGGCAGAAAGCAAGAGAGCTAAAAGCCTCTTGATGA	360
QY	2072	AGTGCAAGAGAGAGTGAATAAGTTGGCTTAAAGCTCAACATTCAGAAAACGGAATCA	2131
DB	361	AAGTGAAGTGGAGAGTGAATAAGTTGGCTTAAAGCTCAACATTCATTAAACTAAGATCA	420
QY	2132	TG 2133	
DB	421	TG 422	

RESULT 42

AA50416/c

ID AAN50416 standard; cDNA; 3649 BP.

XX AC AAN50416;

XX DT 25-MAR-2003 (revised)

DT 08-JAN-1992 (first entry)

XX DE Calf acetyl choline receptor alpha cDNA sequence.

XX KW hACR-alpha; myasthenia gravis; MG; neuromyopathy; ds.

XX OS Homo sapiens.

XX PN JP60078996-A.

PD 04-MAY-1985.

XX PF 05-OCT-1983; 83JP-00186402.

XX PR 05-OCT-1983; 83JP-00186402.

XX PA (MITU) MITSUBISHI CHEM IND LTD.

XX WPI; 1985-144120/24.

XX New peptide(s) useful in treatment of myasthenia gravis - also in

PT diagnosis, having acetylcholine receptor-alpha like activities.

XX Disclosure; Fig 1; 13pp; Japanese.

XX Peptides derived from the hACR-alpha gene product have similar

CC immunogenicity to the gene product, and may be useful in the treatment

CC and diagnosis of myasthenia gravis caused by neuropathy in autoimmune

CC reaction to ACR. See also AAN50415. (Updated on 25-MAR-2003 to correct PA field.)

XX CC

SQ Sequence 3649 BP; 1007 A; 912 C; 780 G; 950 T; 0 U; 0 Other;

Query Match 8.6%; Score 306.2; DB 1; Length 3649;

Best Local Similarity 91.3%; Pred. No. 74;

Matches 335; Conservative 0; Mismatches 28; Indels 4; Gaps 1;

QY 2424 TTTCAGTCTGCTTGGAGAGACTCTTGAGAGTCCCTTGGACTGCAAGGAGATCAA 2483

DB 3000 TTCAGACTGTGCTTGGAGAGACTCTTGAGAGTCCCTTGGACTGCAAGGAGATCAA 2941

QY 2484 CCAGTCCATCTAAAGGAGATCAGTCTCGGCTTTCTTTGGAAGGAATGATGCTAAAGCT 2543

DB 2940 CCAGTCCATCTAAAGGAGATCAGTCTCGGCTTTCTTTGGAAGGAATGATGCTAAAGCT 2881

QY 2544 GAACTCCAGTACTTGGCCACCTGATCAGAGAGCTGACTCTGAAAGACCTGAT 2603

DB 2880 GAACTCCAGTACTTGGCCACCTGATCAGAGAGCTGACTCTGAAAGACCTGAT 2821

QY 2604 GCTGGAGGGATTGGGGCAGGAGAGAGAGGGGACGACAGAGGATGAGTGGCTGGATGG 2663

DB 2820 GCTGGAGGGATTGGGGCAGGAGAGAGAGGGGATGACAGAGGATGAGTGGCTGGATGG 2761

QY 2664 CATCACTGACTGATGACGCTGAGTCTGGGTAAGTCTCTGGAGTTGATGACACAGGA 2723

DB 2760 CATCACTGACTGATGACGCTGAGTCTGGGTAAGTCTCTGGAGTTGATGACACAGGA 2701

QY 2724 GGCCTGTCTCGCGGATTCATGGGTCACAAAGAGTTGGACAGCTGAGCACTGAAC 2783

DB 2700 GGC-----CGTCTGAGTTCTATGGGTCGMAAGAGTCAGACATGACTGAGCGACTGAAT 2645

QY 2784 TGAAGTCT 2790

DB 2644 TGAAGTCT 2638

RESULT 43

ABX46049

ID ABX46049 standard; cDNA; 336 BP.

XX AC ABX46049;

XX DT 21-FEB-2003 (first entry)

XX DE Bovine EST associated with lactation/muscle/fat deposition #11214.

XX KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;

XX KW muscle deposition; fat deposition; genome mapping; gene identification;

XX KW gene analysis; cattle breeding.

XX OS Bos Taurus.

XX FN US2002137139-A1.

XX PD 26-SEP-2002.

XX PF 24-SEP-2001; 2001US-00960352.

XX PR 12-JAN-1999; 99US-0115707P.

XX PR 11-JAN-2000; 2000US-00480902.

XX PA (BYAT/) BYATT J C.

XX PA (MATH/) MATHIALAGAN N.

XX PA (TAON/) TAO N.

XX PA (WARR/) WARREN W C.

XX PI Byatt JC, Mathialagan N, Tao N, Warren WC;

XX DR WPI; 2003-110599/10.

XX PT New nucleic acid associated with lactation, and muscle and fat

PT deposition, useful for genome mapping, gene identification and analysis, cattle breeding, or for genetically improving cattle.

XX Claim 2; SEQ ID NO 11214; 245pp; English.

XX The invention relates to a purified nucleic acid molecule associated with lactation or muscle and fat deposition (designated LMFD) derived from cattle, and the LMFD nucleic acid can specifically hybridize to a second nucleic acid molecule comprising any of 1512 nucleotide sequences, appearing as ABX34836-ABX49947, or complements of them. Also included are (1) a transformed cell having a nucleic acid comprising an LMFD nucleic acid linked to a promoter and a 3' non-translated sequence that functions in the cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and (2) determining a level or pattern of a molecule in a bovine cell or tissue comprising: (a) incubating a marker nucleic acid (comprising any of the 1512 nucleic acid sequences or its complement or fragment) with a complementary nucleic acid molecule obtained from the bovine cell or tissue, where hybridisation between the marker nucleic acid and the complementary nucleic acid permits the detection of the molecule; and (b) detecting the level or pattern of the complementary nucleic acid, where the detection of the complementary nucleic acid is predictive of the level or pattern of the molecule. The LMFD nucleic acid is used for determining a level or pattern of a molecule in a bovine cell or tissue. It is useful for genome mapping, gene identification and analysis, cattle breeding, preparation of constructs for use in cattle gene expression, or for genetically improving cattle. The present sequence is one of the 1512 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The present sequence was not shown in the specification but was obtained in electronic format from the USPTO web site:

XX seqdata.uspto.gov/sequence.html?docID=20020137139

SQ Sequence 336 BP; 88 A; 55 C; 111 G; 81 T; 0 U; 1 Other;

Query Match 8.4%; Score 259.8; DB 1; Length 336;

Best Local Similarity 93.2%; Pred. No. 8.2e+02;

Matches 313; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 2352 GGTTCCTCAGTGTCATGTATGGATGTGAGAGTTGGACTGTGAAGAAAGCTGAGCACTG 2411

DB 1 GGTTCCTCAGTGTCATGTATGGATGTGAGAGTTGGACTGTGAAGAAAGCTGAGTCCG 60

QY 2412 AAGATTGATGCTTTTGAAGTGTGTTGAGAGAGTCTTTGAGAGTCCCTTGAGACTGC 2471

DB 61 AAGATTGATGCTTTTGAAGTGTGTTGAGAGAGTCTTTGAGAGTACCTGAGACTGC 120

QY 2472 AAGGAGATCCCAACAGTCCATTCTGAAGGAGATCAGCCCTGGGATTTCTTTGGAGGAAT 2531

DB 121 AAGGAGATCCCAACAGTCCATTCTGAAGGAGATCAGCCCTGGGATTTCTTTGGAGGAAT 180

QY 2532 GATGCTAAGCTGAACTCCAGTACTTTGGCCACTGATCAGAAAGAGTCACTCACTGGA 2591

DB 181 GATGCTAAGCTGAACTTTCCAGTACTTTGGCCACTCATCGAAGAGTTGACTTATTGGA 240

QY 2592 AAAGACCTGATGCTGGAGGGATTGGGGCAGGAGAGAGGGGACGACAGAGGATGAG 2651

DB 241 AAAGACTGTGATGCTGGAGGGATCCGGNGCAGGAGAGAGGGGACGACAGAGATGAG 300

QY 2652 ATGGCTGATGGCATCACTGACTCGATGGACGTGAG 2687

DB 301 ATGGCTGATGGCATCACTCGACCTGATAGACGTGAG 336

RESULT 44

ABX46780

ID ABX46780 standard; cDNA; 447 BP.

XX AC ABX46780;

XX DT 21-FEB-2003 (first entry)

XX DE Bovine EST associated with lactation/muscle/fat deposition #11945.

KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
 KW muscle deposition; fat deposition; genome mapping; gene identification;
 XX Gene analysis; cattle breeding.
 XX
 OS Bos Taurus.
 XX
 XX US2002137139-A1.
 XX
 XX 26-SEP-2002.
 XX
 XX 24-SEP-2001; 2001US-00960352.
 XX
 XX 12-JAN-1999; 99US-0115707P.
 PR
 PR 11-JAN-2000; 2000US-00480902.
 XX
 XX (BYAT/) BYATT J C.
 PA (MATH/) MATHIALAGAN N.
 PA (TAON/) TAO N.
 PA (WARR/) WARREN W C.
 XX
 XX
 XX Byatt JC, Mathialagan N, Tao N, Warren WC;
 XX WPI; 2003-110599/10.
 XX
 XX New nucleic acid associated with lactation, and muscle and fat
 PT deposition, useful for genome mapping, gene identification and analysis,
 PT cattle breeding, or for genetically improving cattle.
 XX
 XX Claim 2; SEQ ID NO 11945; 245pp; English.
 XX
 XX The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMFD), derived from
 CC cattle, and the LMFD nucleic acid can specifically hybridize to a second
 CC nucleic acid molecule comprising any of 1512 nucleotide sequences,
 CC appearing as ABX34836-ABX49947, or complements of them. Also included are
 CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
 CC acid linked to a promoter and a 3' non- translated sequence that
 CC functions in the cell to cause termination of transcription and addition
 CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
 CC (2) determining a level or pattern of a molecule in a bovine cell or
 CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
 CC of the 1512 nucleic acid sequences or its complement or fragment) with a
 CC complementary nucleic acid molecule obtained from the bovine cell or
 CC tissue, where hybridisation between the marker nucleic acid and the
 CC complementary nucleic acid permits the detection of the molecule; and (b)
 CC detecting the level or pattern of the complementary nucleic acid, where
 CC the detection of the complementary nucleic acid is predictive of the
 CC level or pattern of the molecule. The LMFD nucleic acid is used for
 CC determining a level or pattern of a molecule in a bovine cell or tissue.
 CC It is useful for genome mapping, gene identification and analysis, cattle
 CC breeding, preparation of constructs for use in cattle gene expression, or
 CC for genetically improving cattle. The present sequence is one of the
 CC 1512 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
 CC present sequence was not shown in the specification but was obtained in
 CC electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?DocID=20020137139
 XX
 XX Sequence 447 BP; 162 A; 88 C; 94 G; 103 T; 0 U; 0 Other;
 SQ

Query Match 8.4%; Score 299.4; DB 1; Length 447;
 Best Local Similarity 84.1%; Pred. No. 6.2e+02;
 Matches 360; Conservative 0; Mismatches 66; Indels 2; Gaps 2;
 QY 1048 CAAGACCTAATAGATTGTCACAGAAATGCACTGTCATAGCAACACCTCTTCCAA 1107
 DB 22 CGAAGGCCAGTAAAGTCTTCAAGAGAAATATATTGGTGATAGC-AGAACCCTTTTCAA 80
 QY 1108 CAACACAAGAGAAGACTCTACATCGGACATCCAGATGTCACACCGGAATCAGATT 1167
 DB 81 CGGCGCAAGATRAAGATTCTACATGTTGGACATCATCAGTTAGTCAACACCAAAATCAGATT 140
 QY 1168 GATTATATCTTTTGGAGCAAGATGCGAGAGCTCTATACAGTCAGCAAAAACAGACCA 1227

DB 141 GATTATATCTTTTGGAGCAAAATG-GGAGCTCTATACAGTCAGCAAAAGCAAGCTG 199
 QY 1228 GGAGCTTACTGTGCTCAGATCATGAACCTCTTATTCGCAAAATTCAGACTTAAATTGAAG 1287
 DB 200 GGAGCTGACTGTGCTCTGCTCATGAACACCTATTGTCAAAATACAGACTTAAATTGAAG 259
 QY 1288 AAGTAGGGAACCACTAGATCACTCAGGTAAAGACCTAAATCCATCCCTTTATGATTAT 1347
 DB 260 AAGTAGGGAACCACTAGATCACTCAGGTAAAGACCTAAATCCATCCCTTTACTACTAT 319
 QY 1348 ACAGTGAAGTGAGAAATAGATTAAAGGGCTAGATCTGTAGACAGAGTACCTATATGAA 1407
 DB 320 ACAGTGAAGTGGAATAATATATTTAAGGGACTAGATCTGTAGACAGATGCTGATGAA 379
 QY 1408 CTATGGACAGAGGTTTCATGACATTGTACAGGACAGGATCGAGACCATCCCATCGAA 1467
 DB 380 CTATAGACAGAGGTTTCATGACATTGTACAGGACAGGATCGAGACCATCCCATCGAA 439
 QY 1468 AAGAAATG 1475
 DB 440 GAGAAAG 447

RESULT 45
 AAD26490/c
 ID AAD26490 standard; DNA; 815 BP.
 XX
 AC AAD26490;
 XX
 XX 26-MAR-2002 (first entry)
 DT
 XX Sheep alpha (1, 3) galactosyltransferase post-exon 4 DNA #1.
 DE
 XX Sheep; xenotransplantation; Gal-alpha (1,3)Gal; GAL determinant; surgery;
 KW alpha(1,3) galactosyltransferase; alpha (1, 3) GT; drug screening;
 KW gene therapy; ds.
 OS Ovis sp.
 XX
 XX WO200118096-A2.
 XX
 XX 22-NOV-2001.
 PD
 XX
 XX 14-MAY-2001; 2001WO-US015765.
 PF
 XX
 XX 15-MAY-2000; 2000US-0204148P.
 PR
 XX 13-JUN-2000; 2000US-00593316.
 PR
 XX (GERO-) GERON CORP.
 PA
 XX Denning C, Clark J;
 XX
 XX WPI; 2002-089848/12.
 DR
 XX New ovine tissue devoid of antibody-detectable Gal-alpha(1,3)Gal
 PT determinants, for xenotransplantation, and in the treatment of the human
 PT body by surgery or therapy.
 PT
 XX
 XX Claim 18; Page 76; 86pp; English.
 PS
 XX The patent discloses immunologically compatible animal tissue, suitable
 CC for xenotransplantation into human patients. The invention also relates
 CC to an ovine tissue devoid of antibody-detectable Gal-alpha (1,3)Gal (GAL)
 CC determinants which are made by alpha(1,3) galactosyltransferase (GT). The
 CC ovine tissue is useful for treatment of human body by surgery or therapy
 CC and in xenotransplantation, by transplanting the ovine tissue into a
 CC mammal having circulating antibody against Gal alpha(1,3)GT determinants.
 CC Polynucleotide constructs of the invention are useful for inactivating an
 CC alpha(1,3) GT gene in an ovine cell. Alpha (1,3) GT sequences are useful
 CC for drug screening and for the production of GAL containing synthetic
 CC oligosaccharides. Sequences of the invention are also useful in gene
 CC therapy. The present sequence is alpha (1,3) GT post-exon 4 DNA
 CC

```
SQ Sequence 815 BP; 217 A; 189 C; 186 G; 221 T; 0 U; 2 Other;
Query Match      8.4%; Score 298.6; DB 1; Length 815;
Best Local Similarity 87.8%; Pred. No. 3.4e+02;
Matches 347; Conservative 0; Mismatches 44; Indels 4; Gaps 2;

QY 2402 CTGAGCACTGAAGAATTGATGCTTTTGAACCTGTGTGTGTGGAGAAGACTCTTGAGAGTCC 2461
Db      |||||
QY 815 CTGAGCCCCGAAGAATCGATGCTTTTGAACCTGTGTGTGTGGAGAAGACTCTTGAGAGTCC 756
Db      |||||
QY 2462 CTTGGACTGCAAGGAGATCCACCACTCCATCTGAAGGAGATCAGCCCTGGGATTTCTT 2521
Db      |||||
QY 755 CTTGGACTGCAAGGAGATCCACCACTCCATCTGAAGGAGATCAGCCCTGGGATTTCTT 696
Db      |||||
QY 2522 TGGAGGAATGATGCTAAAGCTGAACCTCCAGTACTTTGGCCACCTGATCAGAAAGACT- 2580
Db      |||||
QY 695 TGGAGGAATGATGCTAAAGCTGAACCTCCAGTACTTAAGGCCACCTCATGAGAAGAAGTT 636
Db      |||||
QY 2581 --GACTCACTGGAAGAAGACCTGATGCTGGGAGGGGATTTGGGGCAGGAGGAGGGGAC 2638
Db      |||||
QY 635 GACTCATATGGAAGAAGACTCATGAAGCTGGGAGGGATTTGGGGCAAGAGGAGAGGGGAC 576
Db      |||||
QY 2639 GACAGAGGATGAGATGGCTGGATGGCATCACTGACTCGATGGAGCTGAGTCTGGGTGAAC 2698
Db      |||||
QY 575 GACAGAGGATGAGATGGCTGGATGGCAACACCACTCAATGGACACGAGTTTGAGTGAAC 516
Db      |||||
QY 2699 TCCTGGAGTTGGTGTGATGGACAGGGAGGGCTGTCTGGCGCGATT- CATGGGTCACAAAG 2757
Db      |||||
QY 515 TCTAGGAGTTGGTGTAGACAGGGAGGGCTGGCGTGTGCGATTACATGGGGTCGCAAG 456
Db      |||||
QY 2758 AGTTGGACACGACTGAGCAACTGAACCTGAAGTGA 2792
Db      |||||
QY 455 AGTMGGACACGACTGAACGACTGAACCTGAAGTGA 421
```

Search completed: August 24, 2004, 13:11:25
Job time : 671 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 13:21:40 ; Search time 604 Seconds
(without alignments)
3.966 Million cell updates/sec

Title: US-10-664-775-2

Perfect score: 3572

Sequence: 1 gtcaggaaggcgcagtg.....gcaacaacagcagaagctt 3572

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 0.5

Searched: 45 seqs, 335272 residues

Total number of hits satisfying chosen parameters: 90

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : rni2.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	672	18.8	3370	1	US-08-060-822A-5
C 2	672	18.8	3370	1	PCT-US94-05357-5
C 3	194.3	5.4	6763	1	US-08-756-506-23
C 4	194.3	5.4	10807	1	US-08-208-176-7
C 5	194.3	5.4	10807	1	US-08-758-506-5
C 6	183.2	5.1	2044	1	US-08-071-601-3
C 7	183.2	5.1	2044	1	US-08-621-100-3
C 8	183.2	5.1	4532	1	US-09-930-377B-1
C 9	164.3	4.6	3317	1	US-09-193-562D-1
C 10	115.6	3.2	19124	1	US-08-487-826B-13
C 11	101	2.8	6669	1	US-10-204-708-6
C 12	100.9	2.8	5862	1	US-10-204-708-63
C 13	100.2	2.8	6070	1	US-10-204-708-10
C 14	100.1	2.8	8961	1	US-10-204-708-80
C 15	97.1003	2.7	5666	1	US-10-204-708-29
C 16	96.1003	2.7	6866	1	US-10-204-708-20
C 17	95.9003	2.7	8607	1	US-10-204-708-72
C 18	95.4003	2.7	6306	1	US-10-204-708-50
C 19	94.7002	2.7	658	1	US-08-998-416-595
C 20	93.9003	2.6	11049	1	US-10-204-708-22
C 21	93.8003	2.6	6326	1	US-10-204-708-57
C 22	92.7003	2.6	10467	1	US-10-204-708-2
C 23	92.4003	2.6	6040	1	US-10-204-708-69
C 24	92.3004	2.6	9347	1	US-10-204-708-35
C 25	91.9999	2.6	7218	1	US-08-232-463-14
C 26	91.5004	2.6	10619	1	US-10-204-708-3
C 27	90.4004	2.5	11049	1	US-10-204-708-23
C 28	89.7004	2.5	8601	1	US-10-204-708-62
C 29	86.9003	2.4	11049	1	US-10-204-708-21
C 30	86.8002	2.4	5455	1	US-10-204-708-33
C 31	85.4004	2.4	636	1	US-08-998-416-1137
C 32	84.5003	2.4	6113	1	US-10-204-708-14
C 33	83.8002	2.3	8093	1	US-10-204-708-31

ALIGNMENTS

RESULT 1

US-08-060-822A-5/c

; Sequence 5, Application US/08060822A

; Patent No. 5432270

; GENERAL INFORMATION:

; APPLICANT: Bevins, Charles L.

; APPLICANT: Diamond, Gill

; APPLICANT: Zasloff, Michael

; TITLE OF INVENTION: No. 5432270el Antimicrobial Peptides

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5432270ris

; STREET: One Liberty Place - 46th Floor

; CITY: Philadelphia

; STATE: Pennsylvania

; COUNTRY: USA

; ZIP: 19103

; COMPUTER READABLE FORM:

; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: PC-DOS

; SOFTWARE: WORDPERFECT 5.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/060,822A

; FILING DATE: 19930511

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/991,200

; FILING DATE: December 15, 1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Rebecca Ralph Gaumond

; REGISTRATION NUMBER: 35,152

; REFERENCE/DOCKET NUMBER: CH-0408

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (215) 568-3100

; TELEFAX: (215) 568-3439

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3370 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; FEATURE:

; NAME/KEY: CAAT signal

; LOCATION: 1382..1389

; FEATURE:

; NAME/KEY: TATA signal

; LOCATION: 1444..1449

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 1503..1560

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 3074..3211

; FEATURE:

NAME/KEY: exon
LOCATION: 1502..1560
FEATURE:
NAME/KEY: exon
LOCATION: 3074..3211
FEATURE:
NAME/KEY: intron
LOCATION: 1..1472
FEATURE:
NAME/KEY: intron
LOCATION: 1561..3073
FEATURE:
NAME/KEY: intron
LOCATION: 3211..3370
US-08-060-822A-5

Query Match 18.8%; Score 672; DB 1; Length 3370;
Best Local Similarity 89.1%; Pred. No. 8.4e-07;
Matches 778; Conservative 0; Mismatches 85; Indels 10; Gaps 5;

QY 1950 TGGGTGGAAGACACAGCTGGAATCAAGATTGCCGGGAGAAATAGCAATAACCTCAG 2009
Db 939 TTGGCTGGAAGACACAGCTGGAATCAAGATTGCCGGGAGAAATGCAATAACCTCAA 880

QY 2010 ATATGAGATGATACACCTTATGCGAGAAAGTGAAGAGAACTAAAAGCCTTTGAT 2069
Db 879 ATATGAGATGATACACCTTATGCGAGAAAGTGAAGAGAACTAAAAGCCTTTGAT 820

QY 2070 GAAGTGAAGAGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTGAGAAAGCAAGAT 2129
Db 819 GAAAGTGAAGATGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTGAGAAAGTGAAG 760

QY 2130 CATGSCATCTGCTCCATCACTTATGCGAAATAGATGGGAAACAGTGAAGAGTGC 2189
Db 759 AATGCGTATGCTCCATCACTTATGCGAAATAGATGGGAAACAGTGAAGAGTGC 700

QY 2190 AGACTTTATTTTGGGGGGCTCAAAATCACTGAGATGCTGACTGACGACATGAAAT 2249
Db 699 AGACTTTATTTTGGGGGGCTCAAAATCACTGAGATGCTGACTGACGACATGAAAT 643

QY 2250 AAAAGACATTTACTCTTGGGAAGAAAGTTA--ACCACTTAGATACATATTGAAAGC 2307
Db 642 AAAACACAGATTACTCTTGGGAAGAAAGTTATGACCAACCTAGATGAT--TTGAAAGC 584

QY 2308 AGAGACATTTACTCTTGGGAAGAAAGTTA--ACCACTTAGATACATATTGAAAGC 2364
Db 583 AGAGATTTACTTTTACCACAAAGGTGCGCTAGTCAAGCCCTATGTTTTTTCAGTG 524

QY 2365 GTCATGATGATGAGTGGAGTTGGACTGTGAAGAAAGCTGAGCACTGAAGATTTGATCT 2424
Db 523 GTCATGATGATGAGTGGAGTTGGACTGTGAAG--AAGCTGAGCGCCAAAGATTGATCT 465

QY 2425 TTTGAACCTGTGTTGGAGAGACTCTTGAAGTCCCTTGGACTGCGCTGAGGAGATCCAAC 2484
Db 464 TTTGAACCTGTGTTGGAGAGACTCTTGAAGTCCCTTGGACTGCGCTGAGGAGATCCAAC 405

QY 2485 CAGTCCATTCTGAAGAGATCAGCCCTGGGATTTCTTTGGAAGGAATGATGCTTAAAGCTG 2544
Db 404 CAGTCCATCTTGAAGAGATCAGTCCCTGGGATTTCTTTGGAAGGAATGATGCTTAAAGCTG 345

QY 2545 AAACCTCAGTACTTTGGCCACCTGATCAGAAAGTGTACTCTGGAAGAAAGCCTTGATG 2604
Db 344 AAGCTCAAATCTTGGCCACCTGATCAGAAAGTGTACTCTTGGAAAGAAAGCCTTGATG 285

QY 2605 CTGGAGGGATTTGGGGCAGGAGAGAGAGGGGACGACAGAGATGAGTGGCTGGATGTC 2664
Db 284 CTGGAGGGATTTGGGGCAGGAGAGAGAGGGGCGACAGAGATGAGTGGTGGATGTC 225

QY 2665 ATCACTGACTGATGAGCTGATGTTGGGTGAACTCTCTGAGTGTGATGAGCAGGGAG 2724
Db 224 ATCAGGACTGATGAGCTGATGTTGGGTGAACTCTCTGAGTGTGATGAGCAGGGAG 165

QY 2725 GCCTGTCTTGGGGGATTCATGGGGTGCACAAAGAGTTGGACACGACTGAGCACTGAACT 2784

Db 164 TCCTGGCATGCTGCAATTCATGGGTGCGAAAGATTGGTACCCTGAGCGACTGAACT 105
QY 2785 GAAGTGAAGTGTACTGAAACCTTAGTAGTTAT 2817
Db 104 GAAGTGAAGTGAAGTGGACTGATGGTTCTCTTT 72

RESULT 2
PCT-US94-05257-5/c
; Sequence 5, Application PC/TUS9405257
; GENERAL INFORMATION:
; APPLICANT: Bevins, Charles L.
; APPLICANT: Diamond, Gill
; APPLICANT: Zaslloff, Michael
; TITLE OF INVENTION: Novel Antimicrobial Peptides
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & Norris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WORDPERFECT 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/05257
; FILING DATE: 11-MAY-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/060,822
; FILING DATE: May 11, 1993
; APPLICATION NUMBER: 07/991,200
; FILING DATE: December 15, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Rebecca Ralph Gaumond
; REGISTRATION NUMBER: 35,152
; REFERENCE/DOCKET NUMBER: CH-0408
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3370 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CAAT signal
; LOCATION: 1382..1389
; FEATURE:
; NAME/KEY: TATA signal
; LOCATION: 1444..1449
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1503..1560
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3074..3211
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1502..1560
; FEATURE:
; NAME/KEY: exon
; LOCATION: 3074..3211
; FEATURE:
; NAME/KEY: intron
; LOCATION: 1..1472

FEATURE:
; NAME/KEY: intron
; LOCATION: 1561...3073
FEATURE:
; NAME/KEY: intron
; LOCATION: 3211...3370
PCT-US94-05257-5

Query Match
Best Local Similarity 89.1%; Score 672; DB 1; Length 3370;
Matches 778; Conservative 0; Mismatches 85; Indels 10; Gaps 5;

Qy 1950 TGGCTGGAAGACACCAAGTGAATCAAGATTGCGGAGAAATAGCAATAACCTCAG 2009
Db 939 TTGCTGGAAGACACCAAGTGAATCAAGATTGCGGAGAAATAGCAATAACCTCAA 880

Qy 2010 ATATGCGAGATATACCAACCTTATGCGAGAAAGTGAAGAGAACTAAAGACCTCTTGAT 2069
Db 879 ATATGCGAGATATACCAACCTTATGCGAGAAAGTGAAGAGAACTAAAGACCTCTTGAT 820

Qy 2070 GAAGTGAAGAGAGAGATGAAAAGTTGGCTTAAAGCTCAACATTCAGAAACGAAGAT 2129
Db 819 GAAAGTGAAGAGAGAGATGAAAAGTTGGCTTAAAGCTCAACATTCAGAAACGAAGAT 760

Qy 2130 CATGCACTCTGCTCCATCACTCATGGAATAGATGGGAAACAGTGAACAGTGC 2189
Db 759 AATGGGTATGATCCATCACTCATGGAATAGATGGGAAACAGTGAACAGTGC 700

Qy 2190 AGACTTTATTTTGGGGGCTCCAAAATCACTGCAGATGGTGAAGTGCAGCCATGAAT 2249
Db 699 AGACTTTATTTTGGGGGCTCCAAAATCACTGCAGATGGTGAAGTGCAGCCATGAAT 643

Qy 2250 AAAAGACACTTACTCTTGGAGAAAGTTA--ACCAACCTAGATAGCATATTGAAAGC 2307
Db 642 AAAACACGATTACTCTTGGAGAAAGTTATGACCAACCTAGATAGCAT--TTGAAAGC 584

Qy 2308 AGAGACATTACTTCCCAACAAAGCCCATCTAGTCAAGGC---TATGGTTTTTCCAGTG 2364
Db 593 AGAGATATTACTTTTACCAACAAAGCTCGCTAGTCAAGGCCTATGTTTTTCCAGTG 524

Qy 2365 GTATGATGATGATGAGATGAGTGTGAGTGTGAAGAAAGCTGAGCACTGAAGATTTGATGCT 2424
Db 523 GTCATGTATGATGATGAGATGAGTGTGAG--AAGCTGAGCGCCAAAGAAATTTGATGCT 465

Qy 2425 TTTGAACGTGGTGTGAGAGAGACTCTTGAGAGTCCCTTGGACTGCAAGGAGATCAAC 2484
Db 464 TTTGAACGTGGTGTGAGAGAGACTCTTGATAGTCTTTTGGACTGCAAGGAGATCAAC 405

Qy 2485 CAGTCCATTCTGAAGGAGATCAGCCCTGGGATTTCTTTTGAAGGAATGATGCTAAAGCTG 2544
Db 404 CAGTCCATCCTAAAGGAGATCAGTCTGGGTGTTCTTTTGAAGGACTCATGCTGAAGCTG 345

Qy 2545 AAATCCAGTACTTTGGCCACTGTATCAGAGAGCTGACTCACTGGAAAGACCTGTATG 2604
Db 344 AAGCTCCAAATACCTTGGCCACTCATGCAAGAGTTGACTCATTTGAAAGACCTGTATG 285

Qy 2605 CTGGAGGAGATTGGGGCAGGAGAGAGAGGAGACACAGAGATGAGATGCTGATGGC 2664
Db 284 CTGGAGGAGATTGGGGCAGGAGAGAGAGGAGGAGGAGAGATGAGATGCTGATGGC 225

Qy 2665 ATCACTGATGATGAGATGATGCTGGTGAACCTCTGGATGCTGGATGATGAGAGAG 2724
Db 224 ATCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 165

Qy 2725 GCTGTCTCTGGGCGATTCTGGGTTCACAAAGAGTTGGACACGACTCAGCAACTGAACT 2784
Db 164 TCTTGGCATGCTGCAATTCATGGGTTCGCAAGAGTTGGTACCACTGAGCGACTGAAT 105

Qy 2785 GAATGAACTGACTGAAACCTTAGTATGTTAT 2817
Db 104 GAATGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACT 72

RESULT 3

US-08-756-506-23
; Sequence 23, Application US/08756506
; Patent No. 5905185
; GENERAL INFORMATION:
; APPLICANT: Garner, Ian
; APPLICANT: Cottingham, Ian R.
; APPLICANT: Temperley, Simon M.
; APPLICANT: Foster, Donald C.
; APPLICANT: Sprecher, Cindy A.
; APPLICANT: Prunkard, Donna E.
; TITLE OF INVENTION: PROTEIN C PRODUCTION IN TRANSGENIC
; ANIMALS
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc. East
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/756,506
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Sawislak, Deborah A
; REGISTRATION NUMBER: 37,438
; REFERENCE/DOCKET NUMBER: 95-28
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6672
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6763 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-756-506-23

Query Match 5.4%; Score 194.3; DB 1; Length 6763;
Best Local Similarity 78.2%; Pred. No. 1.1;
Matches 279; Conservative 0; Mismatches 57; Indels 21; Gaps 4;

Qy 2453 TGAGAGTCCCTTGGACTGCAAGGAGATCCACAGTCCATTTCTGAAGGAGATCAGCCCTG 2512
Db 5268 TGAGAACTGTCTGCACACAGCAGAGTCCACAGTC--ATCCTAAGGAGATCAGTCTG 5325

Qy 2513 GGATTTCTTTGGAAGGAATGATCTAAAGCTGAAACTCCAGTACTTTGGCCACTGATCA 2572
Db 5326 GTGTTCAATT--GGAGGACTGATGTTGAGCTGAAACTCCAACTGTTGGCCACTGATG 5383

Qy 2573 GAAGAGCTGACTCACTGGAAAAACCTGATGTGGAGAGGATTTGGGGCAGGAGAGAA 2632
Db 5384 GAAGAGCTGACTCATTTGAAAAACACCTGATGTGTGGGAAAAGATTGAGGGCAGGAGAGAA 5443

Qy 2633 GGGGACACAGAGATGATGCTGGATGGCATCACTGATCGATGGAGCTGAGTCTGG 2692
Db 5444 GGGGACACAGAGATGATGATGTTGGATGGCATCAACACACATGGACATGGGTTGG 5503

Qy 2693 GTGAACCTCTCTGGAGTTGGTATGGACAGGAGCCCTGTCTGTC-----GGGATTCATG 2746
Db 5504 GTGGACTCCAGGAGTTGGTATGGACAGGAGCCCTGGCTGTACGGAAGCGGTTTATG 5563

Qy 2747 GGGTCACAAAGATTGGACACAGACTGAGCAACTGAACTGAACTGAACTGAACTGAA 2803
Db 5564 GGGTCACAAA-----GACTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 5609

RESULT 4

US-08-206-176-7
; Sequence 7, Application US/08206176
; Patent No. 5639940
; GENERAL INFORMATION:
; APPLICANT: Garner, Ian
; APPLICANT: Dalrymple, Michael A
; APPLICANT: Prunkard, Donna E
; APPLICANT: Foster, Donald C
; TITLE OF INVENTION: Production of Fibrinogen in Transgenic
; TITLE OF INVENTION: Animals
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 4225 Roosevelt Way, N.B.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/206,176
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-648
; REFERENCE/DOCKET NUMBER: 93-15
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-547-8080 ext 322
; TELEFAX: 206-548-2329
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10807 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; IMMEDIATE SOURCES:
; CLONE: ovine beta-lactoglobulin

US-08-206-176-7
Query Match 5.4%; Score 194.3; DB 1; Length 10807;
Best Local Similarity 78.2%; Pred. No. 0.69;
Matches 279; Conservative 0; Mismatches 57; Indels 21; Gaps 4;
QY 2453 TGAGAGTCCCTTGGACTGCAAGGAGATCCAAACAGTCCATTCTGAAGGAGATCAGCCCTG 2512
Db 9312 TGAAACTGTCTGCACACAGCAGAGATCCACCAAGTC--ATCCTAAGGAGATCAGTCTG 9369
QY 2513 GGATTTCTTTGGAGGATGATGCTGAAGCTGAACTCCAGTACTTTGGCCACTGATCA 2572
Db 9370 GTGTTCAIT--GGAGGACTGATGTTGAAGCTGAACTCCAAATGCTTTGGCCACTGATG 9427
QY 2573 GAAGAGTCTCACTCTGAAAGAACCCCTGATGCTGGAGGGATTTGGGGCAGGAGGAA 2632
Db 9428 GAAGAGTCTCACTCTTGAAGAACCCCTGATGCTGGAGAGATTTGGGGCAGGAGGAA 9487
QY 2633 GGGACACAGAGATGAGATGGCTGATGGCATCTGATGCTGATGAGAGTCTGG 2692
Db 9488 GGGACACAGAGATGAGATGGCTGATGGCATCTGATGCTGATGAGAGTCTGG 9547
QY 2693 GTGAATCTCTGGAGTGGTGTGATGACAGGAGGCTGCTCTGC-----GGCGATTTCATG 2746
Db 9548 GTGACTCCAGAGTGGTGTGATGACAGGAGGCTGCTCTGCTACGGAAGCGGTTTATG 9607
QY 2747 GGGTCACAAAGTGTGACACAGTGTGAGCAACTGAACTGAACTGAACTGAACTGAA 2803
Db 9608 GGGTCACAAA-----GACTGAGTGAATGAACTGAACTGAACTGAACTGAA 9653

RESULT 5

US-08-756-506-5
; Sequence 5, Application US/08756506
; Patent No. 5905185
; GENERAL INFORMATION:
; APPLICANT: Garner, Ian
; APPLICANT: Cottingham, Ian R.
; APPLICANT: Temperley, Simon M.
; APPLICANT: Foster, Donald C.
; APPLICANT: Sprecher, Cindy A.
; APPLICANT: Prunkard, Donna E.
; TITLE OF INVENTION: PROTEIN C PRODUCTION IN TRANSGENIC
; TITLE OF INVENTION: ANIMALS
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/756,506
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Sawislak, Deborah A
; REGISTRATION NUMBER: 37,438
; REFERENCE/DOCKET NUMBER: 95-28
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6672
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10807 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-756-506-5
Query Match 5.4%; Score 194.3; DB 1; Length 10807;
Best Local Similarity 78.2%; Pred. No. 0.69;
Matches 279; Conservative 0; Mismatches 57; Indels 21; Gaps 4;
QY 2453 TGAGAGTCCCTTGGACTGCAAGGAGATCCAAACAGTCCATTCTGAAGGAGATCAGCCCTG 2512
Db 9312 TGAAACTGTCTGCACACAGCAGAGATCCACCAAGTC--ATCCTAAGGAGATCAGTCTG 9369
QY 2513 GGATTTCTTTGGAGGAAATGATGCTTAAAGCTGAACTCCAGTACTTTGGCCACTGATCA 2572
Db 9370 GTGTTCAIT--GGAGGACTGATGTTGAAGCTGAACTCCAAATGCTTTGGCCACTGATG 9427
QY 2573 GAAGAGTCTCACTCTGAAAGAACCCCTGATGCTGGAGGGATTTGGGGCAGGAGGAA 2632
Db 9428 GAAGAGTCTCACTCTTGAAGAACCCCTGATGCTGGAGAGATTTGGGGCAGGAGGAA 9487
QY 2633 GGGACACAGAGATGAGATGGCTGATGGCATCTGATGCTGATGAGAGTCTGG 2692
Db 9488 GGGACACAGAGATGAGATGGCTGATGGCATCTGATGCTGATGAGAGTCTGG 9547
QY 2693 GTGAATCTCTGGAGTGGTGTGATGACAGGAGGCTGCTCTGC-----GGCGATTTCATG 2746
Db 9548 GTGACTCCAGAGTGGTGTGATGACAGGAGGCTGCTCTGCTACGGAAGCGGTTTATG 9607
QY 2747 GGGTCACAAAGTGTGACACAGTGTGAGCAACTGAACTGAACTGAACTGAACTGAA 2803

Query Match	5.1%;	Score 183.2;	DB 1;	Length 4532;
Best Local Similarity	81.1%;	Pred. No. 2.3;		
Matches 236;	Conservative 0;	Mismatches 53;	Indels 2;	Gaps 2;
QY	2502	GATCAGCCCTGGGATTCTTCGGAAGGATGATGCTAAAGCTGAAACTCCAGTACTTTGG	2561	
DbB	1	GATCAGTCTCTGGTGTCTATTGAAGGACTGATGCTGGAAGTTGAAGCTCCAACTACTTTGG	60	
QY	2562	CCACCTTGATCAGAAGAGCTGACTCACTGGGAAAGAACCCCTGATCTGGGAGGANTTGGGGG	2621	
DbB	61	CCACCTGATGCGAAGAACTGACTCATGTGATAAGACCCCTGATACTGGGAAGATTGAAGG	120	
QY	2622	CAGGAGGAGAAAGGGACGACAGAGGTAGATGCTGTGGATGGCATCACTGATCTGATGGA	2681	
DbB	121	CAGGAGGAGAA - GGGATGACAGAGGATGGAAGAGTTGGATGGAATCACCAACTCGATGGA	179	
QY	2682	CGTGTAGTCTGGGTGGAAC - TCCTTGGAGTTGGTGATGCACAGGAGGCGCTCTCTCGGGCGCA	2740	
DbB	180	CATGAGTTTGACACAGCTTCCAGGAGTTGGTATATGGGCAGGGAAGCCTGGCTGCTGCAG	239	
QY	2741	TTTCATGGGTCTCAAAGAGTTTGGACACGACTGAGCAACTGAACTGAACTGA	2791	

RESULT 10
US-08-487-836B-13/c
; Sequence 13, Application US/08487826B
; Patent No. 599387
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun


```

; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 80
; LENGTH: 8961
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (5096)
; OTHER INFORMATION: n is a or g or c or t
US-10-204-708-80

Query Match          2.8%; Score 100.1; DB 1; Length 8961;
Best Local Similarity 47.9%; Pred. No. 14; Gaps 2;
Matches 345; Conservative 0; Mismatches 369; Indels 7;

QY 2786 AACTGAACTGTAAGAACTTAGTAGTTATATTAATCTACGAAAAATAGTAATTCATATG 2845
Db 5327 AATAGATATTTTGTGTAATTTTTTTTTTATTTTGGATTTAATGTTAATTTTATAAT 5386

QY 2846 TATTCAAAATATTTTCATAAGTGTGGTTAAGATAAATAAGATTTTCAAATTTGATTTTATC 2905
Db 5387 AATTATATATTTTTTTTTTTTAAGTGTATTTTATTAATAAGTATTTTTCGTAGTATTTAT 5446

QY 2906 TTTGATTTTCTCTACTATTTAAATTTGGGATTTTAACTATTTCTTCAATGACTTGTAT 2965
Db 5447 GTTTATATTTTATAATATTTAGTTTTTATTTTATAGATATTTTATAAATATTTT 5506

QY 2966 TTCTAATATTTACTTATCTATTTTACTTTAATTCGACTTATTTTATTTGATTTTCTAA 3025
Db 5507 TATTTATTTTAGTTTGGATTTCTAATAGTGTAGAAATTTATTTTTTTTATATATAA 5566

QY 3026 TAAATCCAGTCCTTGTTTTTAAAAAGACTTTAAAAATATTA-ATTTCTCTTTAGTGT 3084
Db 5567 TAACGGATATTCGGTTTTTTTTTATATATATATTTTCGTAATAATTTAGTTTTTAT 5626

QY 3085 TTTACAGTCTTTTCAGGCTACTCTTTTGATTTATTTGGTCCATCTTTCTCAAGTTT 3144
Db 5627 ATTGATATTTTATAATAATTAATTTTTTAAAGTTGTTTTTAGTTTGTAGGTTAGGGTT 5686

QY 3145 TGAATTTGGCTACGTAATCATTTATCTTTATTTTTTGTAAATTAGTCTTTAAATTCATT 3204
Db 5687 TAGAGGTTATATATTTTITAGTTTTTTTTTTTTTTTATAGTTTATAGTAAAGTATT 5746

QY 3205 ATTCTTTGATAACAGCTTCAGTTCATAGGCTTTAAATAAGATTTTTTTTTTTTTTTTT 3264
Db 5747 TTTTTTTTATTAATAAAGATAAATTTATAAATGATTAATAGTTTATATGTCGTTTTTATAGATG 5806

```

[illegible]

```
QY 3505 T 3505
Db 6041 T 6041

RESULT 15
US-10-204-708-29
; Sequence 29, Application US/10204708
; Patent No. 667731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 29
; LENGTH: 5666
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chemically treated genomic DNA (Homo sapiens)
US-10-204-708-29

Query Match 2.7%; Score 97.1003; DB 1; Length 5666;
Best Local Similarity 51.5%; Pred. No. 23;
Matches 333; Conservative 0; Mismatches 294; Indels 19; Gaps 5;

QY 2842 TATGATTCACAAATTTATTCATATGTTGGTAAAGATAAAGATTTTCAAAATGATTTT 2301
Db 82 TGGATTTTGAAGAATTTATTTATTTTATTAATTTATTTAGTTTAAAGGTAAGTT 141

QY 2902 TATCTTTGATTTTCTCTACTATTAATTTTGGGATTTTAACTATTTCTTCAATGACCT 2961
Db 142 GGTTTTAGTGTTTTTTTTTATTAATTTTATTTAATGTTGCGATGTTGTTGGTATTT 199

QY 2962 GTATTTCTAATTTACTTATTTCTATTTTACTTTTAACTGACCTATTTTATTTGATTTT 3021
Db 200 TTTATTTTATTAATGAATGCGTGTGTTGTTATTTTATTTTATTTTATTTATTTAGTTT 259

QY 3022 CTAATAAAATCCAGCTCTGTTGTTTTTTA-----AAAAGACTTTTAAAAATTAATTAATTTCT 3075
Db 260 TTAGAGATTTGTTGATTTTATTTGATTTTATTTTAAAGATAAATTTTGTATTTGATTTT 319

QY 3076 CTTTAGTGTTTTACAGCTCTTTTCCAGCTACTCTTTTGTATTTATTTGGTCTATCTTTT 3135
Db 320 TTTATGCTTTTATTTTATTTTAAATTTTATTTGATTTTGTGTTTAAAGTTTATTTGTTTTTT 379

QY 3136 CTCAGTTTGAATTTGGCTAGTAACCAATTTATCTTTATTTTGTGTAATTTAGCTCTTT 3195
Db 380 TTTTGTGTTTTTGGGTTAGTTGTTTATTTTGTATAGATTTTGTGAGCTGATAATTTA 439

QY 3196 AAATCATAATCTTTGATTAACAGCTTCAGTTCTATGCTTTTAAAGATTTT----- 3250
Db 440 GAATATTGAATTCGAAGTTTGTGTTTAAATTTTAAATATGGGTATAAATTTTATTTTATGAT 499

QY 3251 - - - - - TTTTATTTTATTTTAAAGATGTCATCTTTTGTGAAGTTTGTGCAATGCTTTTTCAGCA 3308
Db 500 TGTTTAGTTGATTTTATTAATTTTGAAGTAGAGTTATTTTATTTATTTAGTAGTTAGGAT 559

RESULT 16
US-10-204-708-20
; Sequence 20, Application US/10204708
; Patent No. 667731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 20
; LENGTH: 6866
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chemically treated genomic DNA (Homo sapiens)
US-10-204-708-20

Query Match 2.7%; Score 96.1003; DB 1; Length 6866;
Best Local Similarity 47.4%; Pred. No. 20;
Matches 317; Conservative 0; Mismatches 349; Indels 3; Gaps 1;

QY 2813 TTTATTTACTCAGAAATAGTAATTTTCATATGATTTCAAAATTTATTTCAATATTTGGT 2872
Db 566 TGTGTTTTAAAGAGATTTGTTGTTATTTTATTTAGTGAATTTTTCGCGTATTGGA 625

QY 2873 TAAAGATAATAAGATTTTCAAAATTTGATTTTATTTCTTCTACTTTATTTAA 2929
Db 626 TATATTTAGGGGATTTATTTAGATTTGTTGCGATTAAGTTGTTTATTTTAAATATTTTA 685

QY 2930 TTTTGGGATTTTAACTATTTCTTCAAGACTGTTGATTTCTTAATTTTACTTATTTCTATT 2989
Db 686 GATTTATGCTAGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTAGTT 745

QY 2990 TACTTTAATTTGCACTTATTTTATTTGATTTTCTTAATAAATCCAGCTCTGTTTTTTTA 3049
Db 746 AATTTTGTCTTAATTTTATTTTACGTTAAGTTTATTTTATTTTATTTTATTTTATTTTATTTT 805

QY 3050 AAAAGACTTTTAAATTTAATTTCTTTTGTGTTTATTTTACAGTTCTTTTACAGGCTACTTC 3109
Db 806 AATATAGTTTGGTGAATGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 865

QY 3110 TTTTGTATTTATTTTGTCTCTACTCTTTTCTCAAGTTTTCGAATTTGCTAGTAACCTCATTTAT 3169
Db 866 TAGCTTTTAGTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTATTTAT 925

QY 3309 ATAATTTAGGATATTTTGAATGGTTTCATGATGCTTTTGTACTTGGCATTTATTGAA 3368
Db 560 ATATTTTGAGAAATGTTATTTTAGATAATTTAGATAATTTTATTTTATTTTATTTTATTTTATTTTATTTT 619

QY 3369 GTTTATG-ATTTATGAATTTATGATGCTTTTATTTTGGGCATAAAGTCTTATGCGCATTTT 3427
Db 620 GAGTATATTTTATATAATTTAGATGCTATATTTAATATATTTAGCT---TGGGTTATATG 676

QY 3428 TTGCGTCTATATTTCTTAAATTTAAATTTGGCTTTTAAAGATTT 3473
Db 677 GTATATTTTATTTGTTTATTTAGTTTATAATTTTGTATAGTATAGTT 722
```


QY 2866 TGTTGGTTAAGATAAATAGATTTTCAAATTCATTTTATCTTTGATTTTCTCTACTTAT 2925
DB 5266 TTAAGTTTAAAGTAGGAGAAAGTTTATTAATAATTTTATTAAGATTTTGTATATATC 5325
QY 2926 TTAATTTGGGATTTAACTATTTCTTCAATGACTGTTGATTTCTTAATTTACTTATCT 2985
DB 5326 GTATTTTAAAGGATTTTATTTTAAATGTTTAAAGTGTAGAGTTTGTGTTTATATATA 5385
QY 2986 ATTTTACTTTAATGCACTTATTTTATGATTTTCTTAATAAATCCAGCTCTGTTT 3045
DB 5386 TAGTTCGTTATTTAGATTTTATTTTATGATGTTGTTAGATAGTTTAAATTTTGT 5445
QY 3046 TTTAAAGAGCTTTAAATTTAATTTCTCTTTAGTGTTTTACCAGTTCTTTCAGGCTA 3105
DB 5446 TTTGTTTAAATTTAGTTGTTTAAATGATTTGATTTGATTTTATTTTGTATTTTA 5505
QY 3106 CTTCTT-----TTGATTTATTTGCTCTATCTTTCTCAAGTTTGAATGGC 3153
DB 5506 TTTGTTATTTAAGTTTATAGATATGTTGGAGTTAATTTTATTTTATGATTTATTTT 5565
QY 3154 TACGTAACCTATTTATCTTTATTTTGTAAATAGCTCTTTAAATCAATTTCTTTGA 3213
DB 5566 TTAATTTTGAATTTATTTTATTTTATTTTATAGTGTAGTTTATTAATTTTGTAT -G 5623
QY 3214 TACAGCTTCACTTCTAGGCTTTAAATAGTTTATTTTATTTTATTTTAAAGATG 3273
DB 5624 TTTGGGTTGTGTTTTCGGGAATGTTTGTGTTTATTTTGTGTTTATTTTGTATG 5683
QY 3274 TCATTTCTTTGGAAGTTTGAATGCTTTGAGCAATTAATTTAGGATTTTGAATGGT 3333
DB 5684 AATGAGTTTATGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 5743
QY 3334 TCATGAGTATGCTTTTGTACTGTCATTTATTTGAAGTTTATGATTTATGATTTATGATC 3393
DB 5744 ATTTTGTGATGTTTATTTAGGAAAGTAAATTTTGAATTTTATTTTATTTTATTT 5803
QY 3394 TTTTITTTGGCATAAAGCTATGAGCATATTTTGTGCTATATTTCTTAATTAATA 3453
DB 5804 TATTTTGGTTTATTTTAAATTTTAAAGGTTTATTTTATTTTATTTTATTTTATTT 5863
QY 3454 AATTTGGCTTTAAAGATTTTGTGCTATT 3483
DB 5864 ATTTTATGTTTATGATTTATTTTATGATTT 5893

RESULT 19

US-08-998-416-595
; Sequence 595, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgin
; APPLICANT: Knechtle, Philipp
; APPLICANT: Reibschung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPIL
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 595:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 658 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1408RP
; US-08-998-416-595

Query Match 2.7%; Score 94.7002; DB 1; Length 658;
Best Local Similarity 50.9%; Pred. No. 2e+02;
Matches 248; Conservative 0; Mismatches 239; Indels 1; Gaps 1;

QY 2810 TAGTTTATTTACTCAGAAATAGTAATTTTCATATGATTCAGAAATTTTCAATATTTT 2869
DB 171 TTGTATTAATACCACATTAATCTTTTATCTTGATTAATTAATTTTATTTTGTATTT 230
QY 2870 -GGTTAAGATAAATAGATTTTCAAATTCATTTTATCTTTGATTTTCTCTACTTATTTA 2928
DB 231 ATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 290
QY 2929 ATTTTGGATTTTAACTATTTCTTCAATGACTTGTATTTCTAAATTTTACTTATTTCT 2988
DB 291 ATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 350
QY 2989 TTACTTTTAAATGCACTTATTTTATTTTATTTTATTTTCTAATAAATCCAGTCTTGT 3048
DB 351 TTATTTATTTTGTGTTGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 410
QY 3049 AAAAAGACTTTAAATTAATTAATTTCTCTTTAGTGTGTTTACAGTCTCTTTCCAGGCT 3108
DB 411 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 470
QY 3109 CTTTTGTATTTTGTGCTCTATCTTTTCTCAAGTTTGTGAATTTGGCTAGTAACTCATTTA 3168
DB 471 ATATGTCGACTTTATTTCAATATATAGATTTGTGTTTGTGAACATATATATATGCTATT 530
QY 3169 TCTTTATTTTGTAAATTAGCTCTTTTAAATTCATTTTCTTGAACAGCTTCAAGTTC 3228
DB 531 TCTATTTCTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 590
QY 3229 TATGCTTTTAAATAGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 3288
DB 591 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 650
QY 3289 TTTTGAC 3295
DB 651 TTTTAC 657

RESULT 20

US-10-204-708-22
; Sequence 22, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENROCK, Christian
; APPLICANT: BERLIN, Kurt

RESULT 21
US-10-204-708-57
; Sequence 57, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPNEROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 22
; LENGTH: 11049
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-22

Query Match 2.6%; Score 93.9003; DB 1; Length 11049;
Best Local Similarity 49.1%; Pred. No. 13;
Matches 300; Conservative 0; Mismatches 306; Indels 5; Gaps 2;
2858 TTTCATTAATGTTGGTTAAGATAAAGATTTTCAAATGATTTTATCTTTGATTTTCTAATTTA 2917
1191 TTTCGTAATATGTTAAGATTTTATTTTATTTAGTATTTTATTTAGTTTATTTTATTTT 1250
2918 CTACTTATTTAATTTTGGGATTTTAACTATTTCTCAATGACCTGATTTTCTAATTTA 2977
1251 TTTTGTGTAATTAATTAATTTTATTTTATTTTATATATATTTACGTTTGTATTTATGT 1310
2978 CTTATCTATTTTACTTTAATGACATTTATTTTATGATTTTCTTAATAAATCCAGTC 3037
1311 ATAATGTTGGTTAGACGATTTTGTGATTTTGAATAGAGATTTAGATGAGCTTTTAA 1370
3038 CTTCGTTTTTAAAGACATTTAAATATTAATTTCTCTTTAGTGTGTTTACCAGTTCTT 3097
1371 TTTTATATTTTATAGATTTTAAATTAATGATTAATTAATTAATTTTATGTTTATTTT 1430
3098 TCAGGCTACTTCTTTGATTTATTTTGGTCTATC-TTTTCTCAAGTTTGAATTTGGCTAC 3156
1431 TATTTAGTTGTTATGTTTATTTTATTTTAGTGAATGTTTGTATTTATTTAGTTAGAT 1490
3157 GTAACTCATTATCTTTATTTTGTAAATAGCTCTTTAAATCATATTTCTTTGATAA 3216
1491 GTTGGTATTATTTATATTTATAGTTATTTATAGGTTTATTAATTTTATTTTATGTAAT 1550
3217 CAGCTTCAGTTCTATGCTTTAAATAAGTTTATTTTATTTTATTTTAAAGAAATGCA 3276
1551 TGAGGTATATATAGATTTTATTAACGATTTTGGTGAATTTTAAAGTTTATTTA 1610
3277 TTCTTTGGAAGTTTGTCAATGCTTTGAGCAATTAATTTAGGATATTTTGAATGGTTCA 3336
1611 TTGTTTTTAAGATATATTTTAAATTTTATTTTGGTTAGGAAGTTTGAATGAATG 1670
3337 ---TGAGTATGCTTTTGTACTTGGCATTTATGAGTTTATGATTTATGATTTATGATG 3392
1671 GTTTTGTATTTTGTGTTATTTTATTTTAAATTTTAAATTTTAAATTTATTAAGTTT 1730
3393 CTTTTTTTGGGCATAAAGGCTATGGCATATTTTGTGCTATATTTCTTAAATAT 3452
1731 TTGTTTTTAGGTTTATTTATATGTTTATTTTGTGTTTGAATAGTTTATTTTATTTT 1790
3453 AAATGGCTTT 3463
1791 CGTTTAGTTT 1801

Query Match 2.6%; Score 93.8003; DB 1; Length 6326;
Best Local Similarity 50.4%; Pred. No. 23;
Matches 253; Conservative 0; Mismatches 247; Indels 2; Gaps 1;
2885 ATTTTCAAATGATTTTATCTTTGATTTTCTCTACTATTTAAATTTTGGATTTTAA 2944
2523 ATGTTAGTTTGTGTTAGGTTATTTTGAATTTATTTTATTTTATTTATTTAGGG 2582
2945 TATTTCTTCAATGACCTGTTATTTCTAATATTTTACTTATTTTACTTTTAAATGCACT 3004
2583 TTATTTAGTTTGTGTTAGGTTTATTTTCTGATTTTATTTTATTTATTTAGGG 2642
3005 TATTTTATGATTTTCTAATAAATCCAGCTCTGTTTATTTTAAAGACCTTTAAAT 3064
2643 TTTATTTATTTTATGCTGTTAGTGTAGTGTAGAGAGTTTATTTAATATTTGAAGTTAGAT 2702
3065 TATTAATTTCTTTAGTGTGTTTACCAGTTCTTTCCAGGCTACTTCTTTTGAATTTTGG 3124
2703 ATTTATATTTATGTAATTTTGTGTTTGTGTTTATTTTATTTAGAGTAGATTTATAT 2762
3125 TCTATCTTTTCTCAAGTTTGAATTTGGCTACGTAACCTCATTTATCTTTTATTTTGTAA 3184
2763 TATATTTTGTGTTTATTTTATGATTTTATGAGTTTATTTTGTGTTTATTTTATTTTAA 2822
3185 ATTAGCTCTTTAAATCATATTTCTTTGATAACAGCTTCAGTTCTATGCTTTAATAAAG 3244
2823 GGATTTATTTTGTGTTTATTTTGTGTTTATTTTGTGTTTATTTTATTTAGGTTTATTT 2882
3245 TTTTATTTTATTTTATTTTAAAGAAATGCTATTTCTTGTGAAGTTTTCGCAATGCTTTG 3304
2883 TTTTGTGTTTATTTTGTGTTTATTTATATATATAGTATGAGTTTGGAGGTTAGGATTTG 2942
3305 AGCAATAATTTAGGATATTTTGAATGTTTCAAGATGCTTTTGTCTTCTTGTGCAATTTAT 3364
2943 TGTGTTTGTGTTT-GTATTTTGTGTTTATTTTGGATTAGGTTAGTATGTTAGTTATTTAA 3000
3365 TGAAGTTTATGATTTATGAAT 3386
3001 TAAATAATGGAATATTAAGTT 3022

US-10-204-708-57
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-57

Db	1606	--AATTTATGTTTTTTGTTTTTATTAATGTTATTTTTTTTATTTACGGGATAGAGATTG	1663
Qy	3276	ATTCCTTTGCAAGTTTTGACAAATGCTTTGAGCAATAATTTAGGATATTTTGGAAATGGTTC	3335
Db	1664	TTGTTTTATTTTTAGTTAATGAATTTTTTTTTTAAATTTTTTAAAGT-----ATAGTGGA	1718
Qy	3336	ATGAGTATGCTTTTGTTACTTTGGCAATTTATTGAAGTTTATGCAATTTATGAATATGATGCCTT	3395
Db	1719	GAGAGAATTATTTTTTTTTTTGGGGATATTTTTTAAAGTAGTTTTTTAGTTAGTATATATTTTT	1778
Qy	3396	TTTTTTGGGCATAAAGCTCTATGGCATATTTTTTGT	3431
Db	1779	TTTATTGCTGTTTAAATGATTTTTTTATATATATTTTTTGT	1814

RESULT 24
US-10-204-708-35
Sequence 35, Application US/10204708
Patent No. 6677731
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases associated with DNA Replication
TITLE OF INVENTION: by Assessing DNA Methylation
FILE REFERENCE: 5013.1012
CURRENT APPLICATION NUMBER: US/10/204,708
CURRENT FILING DATE: 2003-05-06
PRIOR APPLICATION NUMBER: PCT/EP01/03971
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: DE 10019058.8
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: DE 10019173.8
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 98
SEQ ID NO 35
LENGTH: 9347
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-35

Query Match	2.6%;	Score 92.3004;	DB 1;	Length 9347;
Best Local Similarity	48.3%;	Prod. No. 16;		
Matches 366;	Conservative	0;	Mismatches 377;	Indels 15; Gaps 4;
QY	2813	TTTATATTACTCAGAAAATAGTAATTCATATGATTCAAAATATTTCATAATGTTGGT	2872	
DB	8555	TTTTTAGTTTGTGGTTATCGTTTTTAGTTTTGTGATTTATTTTAGAGTTTT	8614	
QY	2873	TAAGATAATAAGATTTTCAAAATTCATTTTATCTTTGATTTTTCTCTACATTATTTAATTT	2932	
DB	8615	RAGGGGTTTTAAGATTTAGTTTGGATTTTTTTTTTTTTTTAGATACGTTTATTTTTTT	8674	
QY	2933	TGGGATTTTAACTATTCTTCAATGACCTGTAATTTCTTAATATTTTACTATTCATTTTAC	2992	
DB	8675	GGCGATTTTATTTGGGGTTTTATAGTTTTTATATTTTATGTATTTATTTGATGGTTTTTAAA	8734	
QY	2993	TTTAAATTGCACHTA-----TTTTTATTGATTTTTCTTAATAAAATCCAGTCCTTGTTTTT	3046	
DB	8735	TTTATATTTTCGTTTATTAATTTTTGTTTGAATTTAGTGGTTGTTTGTATTAGTTTGTGTT	8794	
QY	3047	TTAAAAAGACTTTAAAAATTTAAATTTCTTTTAGTTGTTTTACCAGTT--CTTTACGGCT	3104	
DB	8795	TTTAGTTGGAATGTTAAATGGTATTTTAAATTTTGATGTGTTTAAATTAAGTTTTTTCGTA	8854	
QY	3105	ACTCTCTTTTGANTTATTGGGTCCTATCTTTTCTCAAGTTTTGAAATTTGGCTACGTAACCTCA	3164	

[illegible]

RESULT 25
US-08-232-463-14
; Sequence 14, Application US/082322463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29, 768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELE: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:

LENGTH: 7218 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 CLONE: PTZPT-F1s
 US-08-232-463-14

Query Match 2.6%; Score 91.9999; DB 1; Length 7218;
 Best Local Similarity 9.5%; Pred. No. 21;
 Matches 44; Conservative 249; Mismatches 169; Indels 0; Gaps 0;

QY 2840 CATATGTAATCAAAATTTATTCATATGTTGGTAAGATAATAAGATTTTCAAAATGAT 2899
 DB 1012 CATAGCTCACAGATTAATCCGAGCTTGGCTGAGCTCAGGAGCTTCGGAATY 1071
 QY 2900 TTTATCTTTGATTTTCTCTACTTATTAATTTTGGGATTTTAACTATTTCTCAATCAC 2959
 DB 1072 YY 1131
 QY 2960 TTGATTTCTAATATTTACTTATTTACTTTAATTCACACTTATTTTATGATTT 3019
 DB 1132 YY 1191
 QY 3020 TTCTAATAAATCCAGCTCTTCTTTTAAAGAGACTTTAAATATTAATTTCTCTTT 3079
 DB 1192 YY 1251
 QY 3080 AGTGTTTACAGTCTTTCAGGCTACTCTTTGATTTATTTGGTCTCTATCTTCTCA 3139
 DB 1252 YY 1311
 QY 3140 AGTTTGAATGGCTAGCTATCTTATCTTATTTTGTAAATGCTCTTAAAT 3199
 DB 1312 YY 1371
 QY 3200 TCATATCTTTGATTAACAGCTTCAGTCTAGTCTTAAAGTCTTTTCTTTT 3259
 DB 1372 YY 1431
 QY 3260 TTTTAAAGATGCTATCTTCTGAGTTTTCACATGCT 3301
 DB 1432 YYYGTACAAATCTCTATCTCTTAACTACTTGCATAGAT 1473

RESULT 26

US-10-204-708-3
 ; Sequence 3, Application US/10204708
 ; Patent No. 6677731
 ; GENERAL INFORMATION:
 ; APPLICANT: OLEK, Alexander
 ; APPLICANT: PIEPENBROCK, Christian
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
 ; FILE REFERENCE: 5013.1012
 ; CURRENT APPLICATION NUMBER: US/10/204,708
 ; CURRENT FILING DATE: 2003-05-06
 ; PRIOR APPLICATION NUMBER: PCT/EP01/03971
 ; PRIOR FILING DATE: 2001-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019058.8
 ; PRIOR FILING DATE: 2000-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019173.8
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: DE 10032529.7
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: DE 10043826.1
 ; PRIOR FILING DATE: 2000-09-01
 ; NUMBER OF SEQ ID NOS: 98
 ; SEQ ID NO 3
 ; LENGTH: 10619
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 US-10-204-708-3

Query Match 2.6%; Score 91.5004; DB 1; Length 10619;
 Best Local Similarity 47.4%; Pred. No. 15;
 Matches 332; Conservative 0; Mismatches 365; Indels 3; Gaps 2;

QY 2865 ATGTGCTTAAGATAATAAGATTTTCAAAATGATTTTATCTTGAATTTTCTCTACTTA 2924
 DB 3128 ATAGTTGTTAAATTCAGATGATGATTTTGAATTTGTTTAAAGTAA 3187
 QY 2925 TTTAAATTTGGGATTTAACTATTTCTCAATGACTTGTAATTTCTAATAATTTACTTATTC 2984
 DB 3188 AATTTTATTTGTTTATTTATTTATTTATTTATTTATTTTAAATATTTTTT 3247
 QY 2985 TATTTTACTTTTAACTCACTTATTTTATTTATTTATTTTCTAATAAATCCAGCTCTGTTT 3044
 DB 3248 TTTTATTTGAGATAGGTTTATTTATTTAGGTTAGAGTTTAGTGTATTTATG 3307
 QY 3045 TTTTAAAGAGCTTTAAATTTATTTCTCTTTAGTCTTTTACAGTCTTTTCCAGCT 3104
 DB 3308 TTTATTTATAGTTTAAATTTTATTTAGGTTTAGGTTTATTTTATTTTAGTTTTCAGTAG 3367
 QY 3105 ACTTCTTTTGTATTTTGGTCTCTATCTTCTCAAGTTTGAATTTGCTAGTAACTCA 3164
 DB 3368 ATGGGATTAATAGGTTATTTGTTATTTATTTATTTTATTTTGTAGAGATAAGGTTTCTGT 3427
 QY 3165 TTTATCTTTA--TTTCTTAAATTTAGCTCTTTAAATTTTATTTCTTTTGTATAACAGCTT 3222
 DB 3428 ATGTTGTTTACGTTGGTTTGAATTTTGGGTTTAAAGGATTCGTTATTTTAGTTTTT 3487
 QY 3223 CAGTCTCTATGGCTTTAAAGTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 3282
 DB 3488 AAGGTTTAGGATTTAGGTTATGAGTTATTTGTTTATTTTATTTTATTTTAACTTTAAT 3547
 QY 3283 GTCAAGTTTGCACATGCTTTTGAAGCAATTTTGAAGTATTTTGAAGTCTTCAAGTGA 3342
 DB 3548 GGTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 3607
 QY 3343 TGTCTTTGCTTGGCATTATTTGAAGTTTATGATTTATGATTTATGATTTTATTTTATTTTATG 3402
 DB 3608 TGTCTTTTAAAGTATTTAGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 3667
 QY 3403 GGCATAAAGCTCTATGGCATATTTTGTGCTCTATTTTCTTAAATTTTATTTTGGAGTCT 3461
 DB 3668 TGTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 3727
 QY 3462 TTTAAAGTATTTGCTGCTATTTAAACATGAATTTAAGTCTTTATTTTGGACTATAGTGGAGTC 3521
 DB 3728 TTATAATTTTGAATGAATAAATAATGATAGGATTAAGAAAGAAATTAATTTTAAATA 3787
 QY 3522 ACATAAGGTTGGACATGACTTACGCACTAAGCAACACA 3561
 DB 3788 AATTATATTTTGAAGTATAGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 3827

RESULT 27

US-10-204-708-23
 ; Sequence 23, Application US/10204708
 ; Patent No. 6677731
 ; GENERAL INFORMATION:
 ; APPLICANT: OLEK, Alexander
 ; APPLICANT: PIEPENBROCK, Christian
 ; APPLICANT: BERLIN, Kurt
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
 ; FILE REFERENCE: 5013.1012
 ; CURRENT APPLICATION NUMBER: US/10/204,708
 ; CURRENT FILING DATE: 2003-05-06
 ; PRIOR APPLICATION NUMBER: PCT/EP01/03971
 ; PRIOR FILING DATE: 2001-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019058.8

Db 6203 TTTTAAAGGTTGTAATTTTGAAGAAAGATAAAGAAATTTAAGG 6247

RESULT 29

US-10-204-708-21
; Sequence 21, Application US/10204708;
; Patent No. 667731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 21
; LENGTH: 11049
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-21

Query Match 2.4%; Score 86.9003; DB 1; Length 11049;
Best Local Similarity 48.8%; Pred. No. 16;
Matches 313; Conservative 0; Mismatches 321; Indels 7; Gaps 3;

QY 2834 TAATTCATATGTTTCAAAATATTTTCAATAGTGTGTTAAGATAAAGATTTTCAA 2893
Db 7152 TTAATTTGTTTAAATCGGTTTGTATTTTATTTTACGGGAATTTAGATTTGAA 7211
QY 2894 TTGATTTTATCTTGTATTTTCTCTACTTAT-TTAAATTTGGGATTTAACTATTCTT 2952
Db 7212 ATTATTTAGATTTTATTTGTTTGTGTTAATGATTAATTTATTCGTTATTTATTTAGTT 7271
QY 2953 CAATGACTTCTGTTTCTAATATTTTACTTATTTTACTTTTAAATGCACTTTTATA 3012
Db 7272 TTAATATTTTGTGTTTATTTATTTAATTTGTTGTTTAAATTTTATATAGT 7331
QY 3013 TTGATTTTCTAAATAATCCAGTCCTTGTGTTTATTTTAAAGACTTTAAATTTATTAAT 3072
Db 7332 TTAGTGTTTTATTTATATAGTTTAAATGTTGTTGTTAATATGATGAGTAGTTAAT 7391
QY 3073 TCTCTTTAGTGTTTTACCAGTCTTCTTACGCTACTTCTTTGATTTTATTTGTCCTATCT 3132
Db 7392 TTTTGTTTTTATAGATTTTAAATGTTTGTGTAATATGCGTTTAAAGTGTATAGATAT 7451
QY 3133 TTTCTCAAGTTTGAATTTGGCTACGTAACCTATTATCTTTTATTTTCTGTAATTAGCTC 3192
Db 7452 ATTAATTTTATAGTTATGATTAATTAATCGAGATTTTAAATTTTAAATTAAGATTA 7511
QY 3193 TTTAAATTCATATTTCTTTCATACAGCTTCAGTCTATGCTTGTGTTTAAATTAAGTTT 3252
Db 7512 AAGAAATTTATTTTCTGTTATGAAATGTTATTTGTTTGGTTTATATCGAAAGTAT 7571
QY 3253 TTTTATTTTATTTTAAAGATGTCATCTTT-----GTGAAGTTTTCACAATGCTTTGAGC 3307
Db 7572 TTTTGGTTTGTGTTAGATTTTAAATTTTGTTCGTAGAAATTTAAATTTTCCGTATATA 7631
QY 3308 AATAATTTAGGATTTTGTGAATGTTTCATGAGTAGTCTTTTGTACT-TGGCATTTATTTG 3366

Db 7632 AATTTATGAAATTAGATAAGTTAAATTTATATAAATATTTTAGATTGTTTATTTT 7691
QY 3367 AAGTTTATGATTTATGAATTTATGATCTTTTTCGGCATAAAGTCTATGCGATATTT 3426
Db 7692 TATTTTGTGTTTAAATTTATTTATGTAATTTTATAGTCGTTTAAATTTTATATAT 7751
QY 3427 TTTGTCGCTATATTTCTTAAATTTATAAATTTGCGTTTAAAA 3467
Db 7752 TTTTAAAGATTTTAAAGTTAAATTTTAAATTTAGTTGAGAA 7792

RESULT 30

US-10-204-708-33
; Sequence 33, Application US/10204708
; Patent No. 667731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 33
; LENGTH: 5455
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-33

Query Match 2.4%; Score 86.8002; DB 1; Length 5455;
Best Local Similarity 48.5%; Pred. No. 32;
Matches 344; Conservative 0; Mismatches 347; Indels 18; Gaps 4;

QY 2806 TTAGTAGTTTATATTACTCAGAAAATAGTAATTTTATATGTTTCAAAATTTATTTCTATA 2865
Db 430 TTTGTTATTTAAATGTTATAGGATTTTATTTTATTTATTTATTTTATTTTATTTTATTT 489
QY 2866 TGTGTTTAAATTAAGATTTTCAATTTGATTTTATCTTGAATTTTCTCTACTTAT 2925
Db 490 AGTTGGGTAAGTTTATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 546
QY 2926 TTAATTTTGGGATTTTAACTATTCTTCAATGACTTCTTATTTCTAATATTACTTATCT 2985
Db 547 TTTTATTTTTCGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 606
QY 2986 ATTTTACTTTAATTTGCACTTATTTTATTTGATTTTCTAATAAATCCAGTCTTGTGTTT 3045
Db 607 GTTTTGTGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 663
QY 3046 TTTAAAAAGACTTTAAATTTTAAATTTCTCTTTAGTGTTTTACCAGTTCTTTTACGCTA 3105
Db 664 ATTCGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 723
QY 3106 CTCCTTTTGAATTTATTTGGTCTTATCTTTCTCAAGTTTGAATTTGGTACGTAACCTCAT 3165
Db 724 GTTTTATTAAGTTTATTTTAAATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTAT 783
QY 3166 TTATCTTTATTTTGTGTAATTTAGTCTTTTAAATTTCAATTTATTTTGTGATAACAGCTTCAG 3225


```

; TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/867,106
; FILING DATE: 19920625
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PJ 7187
; APPLICATION NUMBER: PCT/AU90/00530
; FILING DATE: 02-NOV-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Feeney, Joanne Longo
; REGISTRATION NUMBER: 35,134
; REFERENCE/DOCKET NUMBER: RICE-0002
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5852 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2378..5038
; NAME/KEY: CDS
; LOCATION: 2378..5038
; US-07-867-106-2

; Query Match 2.3%; Score 82.6003; DB 1; Length 5852;
; Best Local Similarity 46.1%; Pred. No. 33; Mismatches 369; Indels 22; Gaps 2;
; Matches 335; Conservative 0;

QY 2796 TACTGAAACCTTAGTAGTTATATCTACGAAATAGTAATTTTCATATGATATTCATAAT 2855
DB 1772 TATCGATATACITTAATTTATTAAGATGGAATATATTTTAAATTTTAAATAATCTAC 1831
QY 2856 TATTCATATAGTTGGTTAAGATAATAGATTTTCAAAATGATTTTATCTTTGATTTT 2915
DB 1832 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1891
QY 2916 CTCTACTATTATTAATTTTGGGATTTTAACTATTCT-----TCAATGAC 2959
DB 1892 TAATTAATAATCAATTTAAATTAAGTAATTTTAAATATGCAAAACTATAAAAAAC 1951
QY 2960 TTGTATTTTCHAAATTTACTATTCTATTATTTACCTTTTAAATGCACTTATTTTATGATT 3019
DB 1952 TAAAGTAGTTTAAACCAACTTTTCTATTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2011
QY 3020 TTCTAATAAATCCAGCTCTTTGTTTTTAAAGACCTTTAAATTAATTAATTTCTCTTT 3079
DB 2012 GAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2071
QY 3080 AGTGTGTTTACCAGTTCTTTTCAGGCTACTCTTTTGAATTTATTTGGTCTCTATCTTTTCTCA 3139
DB 2072 TTTTCTTTTGAATTTTCTTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2131
QY 3140 AGTTTGAATGGCTACGTAAGTCTATTTATCTTTATTTTCTTTTGAATAGCTCTTTAAAT 3199

; Db 2132 TTTATAAATATTGGATAATATATCAAAATATTTATCAGTTTGGCATGACAAATTTAATTA 2191
; QY 3200 TCATTATTTCTTTGATAACAGCTTCAGTTCTATGGCTTTAATAAAGTTTTTTTTTTTTTTT 3259
; Db 2192 TATTTATTTTGTATTAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2251
; QY 3260 TTTTAAAGAAATGTCATTTCTTTGGAAGTTTTCACATGCTTTGACCAATTAATTAGGA 3319
; Db 2252 TTAATTTTAAATTTTAAATTTTATTTTCCACACTTTTCATTTTATTTTATTTTATTTAT 2311
; QY 3320 TATTTTGAATGGTTTCATGAGTATGCTTTTCTACTTGGCATTATTTGAAGTTTATGATTT 3379
; Db 2312 TGTAAATTCATTTATTTTATTTTAAATAGATTTTGGTTTAAATTTTATTTCAAGATT 2371
; QY 3380 ATGAATATGATGCTTTTCTTTTGGGCATAAAGGCTCTATGGCATATTTTGTGGTCTATA 3439
; Db 2372 TTAATAATGGACAACTTATTTCTTTGGGATAGGTTTTTAAAGTTTTTTGTA-----ATA 2425
; QY 3440 TTTCTAATAATTAATAATTTGGCTTTAAAGAGTATTTGCTGCTATTAAACATGAATTAAGTC 3499
; Db 2426 CTTTGGGAAGAAATCCAAAGTTTGAAGAAATGATGCGTTTGCAGTGTGCAATTATGAC 2485
; QY 3500 TTATTT 3505
; Db 2486 ATTCTT 2491

; RESULT 35
; US-10-204-708-86
; Sequence 86, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: FIEFENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: By Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 86
; TYPE: DNA
; LENGTH: 11050
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; US-10-204-708-86

; Query Match 2.3%; Score 82.4003; DB 1; Length 11050;
; Best Local Similarity 46.5%; Pred. No. 18; Mismatches 306; Indels 0; Gaps 0;
; Matches 266; Conservative 0;

QY 2923 TATTTAATTTTGGGATTTTAACTATTCTTCAATGACTTGTATTTCTAATAATTTACTTAT 2982
DB 7114 TTTTATGTTTGGTTTCTTTTAAAGTCGATTTTAAAGTATTTTATTTAGATTGAGTTT 7173
QY 2983 TCTATTTTACTTTAATTTGCACTTATTTTATTTATTTCTAATAAATCCAGCTCTCTCT 3042
DB 7174 TATATTTTAAATTTATTTATGAGTTTATTTATTTAGTTTAAAGAAATGAGTTAATTTTA 7233
QY 3043 TTTTAAAAAGAACTTTAAATTTATTTCTTTTAGTGTTTTACCAGTTCTTTTCAGG 3102
```

```

Db 7234 TTAATTTTATTTAAGGATTTTCGAGGAAATTAATCGGTATTATAGATTCGGTTTCGGA 7293
QY 3103 CTACTCTCTTTGATTTATTTGGTCTCTATCTTTCTCAAGTTTTCGAAATGGCTAGCTAACT 3162
Db 7294 TTTAAGGATTTTATTTTGTAAATCGTTAGGTGTTTATGATTTGAGTTAGTTTCGTAAT 7353
QY 3163 CATTTATCTTTATTTTGTGAAATAGCTCTTTAAATCATTATCTTTGATTAACAGCTT 3222
Db 7354 TTGTGTTTCGGGGTTTAAATATTAGTTTGGGTTTTTTTATATTATTTTATGTTTATTTA 7413
QY 3223 CAGTTCTATGCTTTAATAAGATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 3282
Db 7414 GATTTTGAATTTAGTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 7473
QY 3283 GTGAAGTTTTCGCAATGCTTTGAGCAATAATTAGGATAATTTTGAATGTTTCATGAGTA 3342
Db 7474 GAGGATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7533
QY 3343 TGCCTTTTGTACTTCGCAATTTATTTGAAGTTTATGATTTATGATGCTTTTCTTTTCTTTT 3402
Db 7534 TTTTCTTTAGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 7593
QY 3403 GGCATAAGGTTCTATGGCAATTTTTTGTGCTCTATATCTTTAAATTAATAATTTGGCTT 3462
Db 7594 AAGTTAGTAGTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 7653
QY 3463 TAAAAAGTATTTGCTGCTATTAACATGAAT 3494
Db 7654 TATTTTTTTTGGAGTTAGTAGATTTTGTGT 7685

RESULT 36
US-10-204-708-55
; Sequence 55, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 55
; LENGTH: 11015
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-55

Query Match 2.3%; Score 81.3004; DB 1; Length 11015;
Best Local Similarity 46.7%; Pred. No. 19;
Matches 316; Conservative 0; Mismatches 347; Indels 13; Gaps 2;

QY 2813 TTTATATTACTAGAAAATAGTAAATTCATATGATATTCATTAATTTTCTATGTTGT 2872
Db 9215 TATAGTTAGTGAATTTGTTTATTAATAATGCTGTTATTTATTTATTTATTTAGTTTGA 9274
QY 2873 TAAGATAATAAGATTTTCAAAATGATTTTATCTTT-----GATTTTCTCTACTTATTT 2927

```

```

Db 9275 AGGTTTTTTTTTTTAAAGAGGAAATTTTATATTTTAAAGTAGTATTTTATTTTATTT 9334
QY 2928 AATTTTGGGATTTTAACTATTTCTTCAATGACTGCTGATTTCTTAATTTTAAATTTACTTATCTAT 2987
Db 9335 TTTTCTTTAGTTTATTTTAAATTAAGAAATTTGTTTTTTTTTTTATCGAATTTGTTTCTTTG 9394
QY 2988 TTTACTTTAAATTCGACTTATTTTATTTGATTTTCTTAATAAAATCCAGTCCCTGTTTTTT 3047
Db 9395 GGTATTTTATAAGTCGATTTATATTTTATATGTTTTTTTGTGTAGGTTTTTTTTTATTT 9454
QY 3048 TAAAGAAGCTTTTAAATTAATTTCTTTTCTTTAGTTTCTTACCAGTTCTTTCAGGCTACT 3107
Db 9455 TAATATAATTTTGAAGTTTATTTATTTGTTGATTAATGATTAATATATTTTATTTTATTTT 9514
QY 3108 TCTTTTGAATTTATTTGGTCTCTATCTTTTCTCAAGTT-----TTGAATTTGGCTACGTA 3159
Db 9515 TTTTCTTTGTTTTTTCGAGAAGGAGTTTTTTTTTATTTGCTGTTTAGTTGGAGTGTAGTGATG 9574
QY 3160 ACTCAITTTATCTTTATTTTCTTTGTAATTTAGCTCTTTAAATTCATTTCTTTGATAACAG 3219
Db 9575 AGATTCGGTTTATTTGTAATTTATCTGTTTTTTTAGGTTTAAAGCAATTTTCTTTTCTTTT 9634
QY 3220 CTTTCAGTTCTATGGCTTTTAAATAAGTTTTTTTTTTTTTTTTTTTTTTTTTAAAGATGTCATTC 3279
Db 9635 TTTAAGTAGTTGGGATTTATAGGATTTTATTTATTTAGCTTTAGTTAAATTTTGTATTTTA 9694
QY 3280 TTTGTGAAGTTTGTACAATGCTTTGAGCAATAATTTAGGATAATTTTGAATTTGGTTTCATGA 3339
Db 9695 GTAGAGATAGGGTTTTTATTTATTTGTTAGGTTGTTTTTAAATTTTTCATTTTCGTGATTC 9754
QY 3340 GTATGCTTTTGTACTTGGCATTATTTGAAGTTTATGATTTATGATTTATGATGCTTTTTT 3399
Db 9755 GTTATTTTATTTTAAAGTTTGGAAATTTATAGGTATTTAGTTAGTTATTTTGTATTTTA 9814
QY 3400 TTGGGCATAAAGTCTATGGCATAATTTTGTGCTCTATATTTCTTAAATTTATAAATGG 3459
Db 9815 TTATTTTTTTTATGTTGTAATAATATTTTATTTATATGATAGATTAATTTTATTTTAT 9874
QY 3460 CTTTAAAAAGTATTTG 3475
Db 9875 TATATATTAGTTGATG 9890

RESULT 37
US-10-204-708-45
; Sequence 45, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 45
; LENGTH: 19233
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

```

OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
FEATURE:
NAME/KEY: unsure
LOCATION: (76, 178..179, 273, 586, 648, 651, 920, 1014, 1173, 1197, 1228)
OTHER INFORMATION: n is a or g or c or t
FEATURE:
NAME/KEY: unsure
LOCATION: (1264, 1267, 1559, 1607, 1677, 1703, 1742, 1848, 1850, 2033)
OTHER INFORMATION: n is a or g or c or t
FEATURE:
NAME/KEY: unsure
LOCATION: (2168, 2254, 2257, 2280, 2291, 2723, 3018, 3080, 4205, 4209)
OTHER INFORMATION: n is a or g or c or t
FEATURE:
NAME/KEY: unsure
LOCATION: (4216, 4225, 4239, 4246, 4254, 4265, 4270, 4274, 4288, 4294..4295)
OTHER INFORMATION: n is a or g or c or t
FEATURE:
NAME/KEY: unsure
LOCATION: (4298, 4363, 4610, 4627, 4675, 4693, 4698, 4738, 4777, 4781)
OTHER INFORMATION: n is a or g or c or t
FEATURE:
NAME/KEY: unsure
LOCATION: (4787, 4825, 4829, 4838, 4855, 5462, 5494, 5496, 5500, 5674)
OTHER INFORMATION: n is a or g or c or t
FEATURE:
NAME/KEY: unsure
LOCATION: (5690, 5697, 5705, 5715, 5890, 5910, 5913, 6027, 6080, 6109)
OTHER INFORMATION: n is a or g or c or t
FEATURE:
NAME/KEY: unsure
LOCATION: (6244, 6288, 6776, 6873, 6891, 6980, 7028, 7656, 7876, 8030)
OTHER INFORMATION: n is a or g or c or t
FEATURE:
NAME/KEY: unsure
LOCATION: (8059, 8075, 8632, 8636, 9700, 9717, 9815, 9818, 9867, 9879)
OTHER INFORMATION: n is a or g or c or t
FEATURE:
NAME/KEY: unsure
LOCATION: (10005, 10075, 10150, 10159, 10165, 10356, 11123, 11166, 11360)
OTHER INFORMATION: n is a or g or c or t
FEATURE:
NAME/KEY: unsure
LOCATION: (11703, 11933, 12050, 12340, 12981, 13564, 13764..13765, 13787)
OTHER INFORMATION: n is a or g or c or t
FEATURE:
NAME/KEY: unsure
LOCATION: (13846, 14168, 16566, 17006, 17217, 17650, 17653, 17656, 17898)
OTHER INFORMATION: n is a or g or c or t
FEATURE:
NAME/KEY: unsure
LOCATION: (18130, 18143, 18151, 18188, 18498, 18549, 18936, 18946, 19165)
OTHER INFORMATION: n is a or g or c or t
FEATURE:
NAME/KEY: unsure
LOCATION: (19200)
OTHER INFORMATION: n is a or g or c or t
US-10-204-708-45

Query Match 2.3%; Score 81.3003; DB 1; Length 19233;
Best Local Similarity 48.1%; Pred. No. 11;
Matches 326; Conservative 0; Mismatches 335; Indels 17; Gaps 4;
QY 2804 CCTAGTAGTTTATATTAATCTACTCAGAAATAGTAATTCATATGTAATCCAAATATTTTCAT 2863
DB 5799 CGTTATTATATTCGGTTATTTTGTGTTTATTAGAGATTGGGTTTATTATGTTGGT 5858
QY 2864 AATGTCGTTAAGATAAATTAAGATTTCAAATGATTTTATCTTTGATTTTCTCTACTT 2923
DB 5859 TAGGTGATTCGAAATATAGATTTAAGTTNATTCGTTTCGTTTNTTNAAT 5918
QY 2924 ATTTAAATTCGGATTTTAACTATTCCTCAATGACTGTATTTCTAATATTTACTTATT 2983

DB 5919 GTTGGAAATTATAGGTATGAATT-----ATCGCGTTTGGTTTAAAGTAATTTTATT 5969
QY 2984 CTATTTTACTTAAATGACACTATTTTATGATTTTCTAATAAAATCCAGTCCTGTT 3043
DB 5970 GTTATTTTAAATATATATGCTAGTCTTTTATAAATTTTGTATAGAGTTATTTTANT 6029
QY 3044 TTTTAAAAAGACTTTAAAAATTAATTTCTCTTTAGTGTTTTACCAGTCTTTTCAGGC 3103
DB 6030 TTAAGAATAAATTTTCGTTATTTCTGTTTATTTATGTTTAGAGGNTTTAAGTT 6089
QY 3104 TACTTCTTTTGATTA---TTTGGTCCATCTTTTCTCAAGTTTGAATTTGGCTACGTA 3160
DB 6090 TAGTATACGTATTTTAAAGNTAAATTTTATTTTAAAGTTTGGGTTTGTGATTTT 6149
QY 3161 CTCATTATCTTTATTTTGTGAAATAGCTTTTAAATTCATTTCTTTGATACAGC 3220
DB 6150 AITTTTACGATGATTAATTAATTTATTTAGTTAGTGTAAATGAAATTTGAGAAATTA 6209
QY 3221 TTCAGTTCATGGCTTTTAAAGTTTATTTTATTTTATTTTATTTTAAAGATGTCAT--- 3277
DB 6210 TTTTATTTATATGTTTATTTATTTGGTATTTAGNTTTTACGAAGTGAAGTATTTT 6269
QY 3278 -TCTTTGCAAGTTTGTACAAATGCTTTGAGCAATAATTTAGGATATTTTGAATGTTCA 3336
DB 6270 GTTTTGGGTAGATTAATTAAGTTTATTTGGTATTTTATTTATTTATTTATTTTATTT 6329
QY 3337 TGAGTATGCTTTTGTACTTGGCATTTATGAAAGTTTATGATTTATGATTTATGATGCTTT 3396
DB 6330 TTTTGTGTTGAGTTGTTGATTTTATTTAGTATACGTTTGTATTTATTTTATTTTATTT 6388
QY 3397 TTTTGGGCAATAAGGCTCTATGCGCATTTTGTGGTCTATATTTCTTAAATTTATAAT 3456
DB 6389 ATAGTATATTTTATGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 6448
QY 3457 TGGCTTTTAAAAAGTATTT 3474
DB 6449 TTGGTATATTTTATTT 6466

RESULT 38
US-08-213-419B-3/c
; Sequence 3, Application US/08213419B
; Patent No. 6333406
; GENERAL INFORMATION:
; APPLICANT: Inselburg, J. et al.
; TITLE OF INVENTION: GENE ENCODING PROTEIN ANTIGENS OF PLASMODIUM FALCIPARUM
; FILE REFERENCE: JII-002CNC
; CURRENT APPLICATION NUMBER: US/08/213,419B
; PRIOR FILING DATE: 1994-03-14
; PRIOR FILING DATE: 1992-04-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 6124
; TYPE: DNA
; ORGANISM: Plasmodium falciparum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2407)..(2439)
; NAME/KEY: CDS
; LOCATION: (3259)..(3404)
; NAME/KEY: CDS
; LOCATION: (3580)..(3720)
; NAME/KEY: CDS
; LOCATION: (3850)..(5835)
US-08-213-419B-3

Query Match 2.3%; Score 81.0002; DB 1; Length 6124;
Best Local Similarity 47.2%; Pred. No. 33;
Matches 246; Conservative 0; Mismatches 275; Indels 0; Gaps 0;


```

; SEQ ID NO 41
; LENGTH: 8537
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-41

Query Match      2.2%; Score 80.1003; DB 1; Length 8537;
Best Local Similarity 47.1%; Pred. No. 25;
Matches 335; Conservative 0; Mismatches 369; Indels 7; Gaps 3;

Qy 2799 TGAACCTTAGTATGATATATCTACAGAAAGTAGTAATTCATATGATTCATAATCAAAATTAAT 2858
Db      |||||
Qy 6229 TGTAGGTTAGTATTTAGGAGGTTGAGAGGGGAGATTTAGTTGGAGGTATAGTTG 6288
Db      |||||
Qy 2859 TCAATAAGTGGTTAGATA----ATAAGATTTCAAATGATTTTATCTTTGATTTT 2914
Db      |||||
Qy 6289 TAGTGAGTTGATTAGTTATTTGATTTTAAATTTGGCGATAGAGTAAGATAAAATGAA 6348
Db      |||||
Qy 2915 TCTCTACTTATTAATTTTGGGATTTTAACTATTTCTCAATGACTTGTATTTCTAATAT 2974
Db      |||||
Qy 6349 TATAATTTAATTAATTAATTAATTAATGAATGGTTTGAAGATTTTGTGTAGTAT 6408
Db      |||||
Qy 2975 TTACTTATCTATTTTACTTTAATGCACTTATTTTATGATTTTCTTAATAAAATCCA 3034
Db      |||||
Qy 6409 TTAATATTTTATTTTATAGTGTAAATTTGATATATTTGTTTAAATGAGATGAATTA 6468
Db      |||||
Qy 3035 GTCCCTGTTTTTAAAGACITTAATATTAATTTCTCTTAGTGTTTTACCAGTT 3094
Db      |||||
Qy 6469 ATGTTGTTGTTTGAATTTTATATAAAATATATTAGGAGTGTGAGTGTTCGTTGT 6528
Db      |||||
Qy 3095 CTTTCAGGCTACTCTTTGATTTATTTGGTCTATCTTTCTCAAGTTTGAATGGCT 3154
Db      |||||
Qy 6529 TTATATGTTTGTAAACAGATATTTTGTAGTTTGTGATTTTAAATTTTGTGTT 6588
Db      |||||
Qy 3155 ACGTA--ACTCATTTATCTTTATTTTGTAAATGACCTCTTAAATCAATATCTTTG 3212
Db      |||||
Qy 6589 ATATAGTAAGTTTAAATGATTTTAAATTTTGTGATTAATAATGATGTGGATATTTG 6648
Db      |||||
Qy 3213 ATAACAGCTTCAGTCTATGCTTTAAATAAGTTTTTTTTTTTTTTTTTTTTTAAAGAA 3272
Db      |||||
Qy 6649 TTTATATGTTTGTATTTTGAATTTTGTGATGTTGTTGATGTTTAAATTTTGTGTT 6708
Db      |||||
Qy 3273 GTCAATCTTGTGAAGTTTGAACAATGCTTTGACCAATAATTTAGGATATTTTGAATGG 3332
Db      |||||
Qy 6709 ATTTT-TTATATTTTATTTTATGATGTAAGTTTATATATATTTGTTATAGTTT 6767
Db      |||||
Qy 3333 TTCAATGATATGCTTTTCTACTTGGCATTTATGAGCTTTATGATTTATGAATATGATG 3392
Db      |||||
Qy 6768 TTAGTATGTTTGTGTTAAATGTTTTTTTATGTTGTTGATTTGTTATTTTATTAGTG 6827
Db      |||||
Qy 3393 CTTTTTTTGGCATAAAGGCTATGCGATATTTTGTGCTATATTTCTTAAATATAT 3452
Db      |||||
Qy 6828 GTGTTTTTGTATGATAGAGTGTGAATTTATTTTATTTATTTATTTATTTATTTAT 6887
Db      |||||
Qy 3453 AAATGGCTTTAAAGATTTTGTGCTATTTAAACATGAATTAAGTCTTAT 3503
Db      |||||
Qy 6888 TTTTTTTTTGAGATAGAGTTTGTGTTTTTGTGTTTGTAGTTGGAGTGAAT 6938
Db      |||||

RESULT 41
US-08-998-416-186
; Sequence 186, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Reischung, Corinne

```

```

; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPHII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSES: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 186:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 615 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1074RP
; US-08-998-416-186

Query Match      2.2%; Score 79.7003; DB 1; Length 615;
Best Local Similarity 49.6%; Pred. No. 3.3e+02;
Matches 280; Conservative 0; Mismatches 278; Indels 7; Gaps 3;

Qy 2843 ATGTATTCAAAATATTTCATTAATGTTGGTTAAGATAATAAGATTTTCAAAATGATTTT 2902
Db      |||||
Qy 2903 ATCTTTGATTTTCTCTACTTATTTAAATTTGGGATTTTAACTATTTCTTCAATGACTTG 2962
Db      |||||
Qy 82 TATTATATCATTTTATTAATAATTAATTTATTTGATTTATTAATATATATATAT 141
Qy 2963 TATTTCTAATATTTACTTATTTCTATTTTACTTTAATTTGCACCTATTTTATTTGATTTTC 3022
Db      |||||
Qy 142 TATATAATTTACTTAATTCATCATTTAATTAATTTATATATATATAATAATAATAATTTAA 201
Qy 3023 T----AATAAAATCCAGTCCCTGTTT---TTTTAAAGAGACTTTAAATATTAATTTCTC 3076
Db      |||||
Qy 202 TATGAATATCTATTATGCTATGTTCAAAATTTTAAATAGTTTAAATATTAATTTATAGATA 261
Qy 3077 TTTAGTGTGTTTACCAGTTCTTTTCAGGCTACTTCTTTTGATTTTATTTGGTCTATCTTTTC 3136
Db      |||||
Qy 262 TTTATTTTCTTTTAAATATTAATAATTAATAATTAATAATAATAATAATAATAATAATA 321
Qy 3137 TCAAGTTTGAATTTGGCTAGTAACATCATTTATCTTTATTTTGTAAATTAGCTCTTTA 3196
Db      |||||
Qy 322 ATGTTTATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 381
Qy 3197 AATTCATTTCTTTGATAACAGCTTCAGTTCTATGCTGCTTAAATAAGTTTTTTTTTTT 3256
Db      |||||
Qy 382 AATTATATTTTATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 441
Qy 3257 TTTTTTTTTTAAAGATGTCATCTTTGTTGAAAGTTTTCGACAAATGCTTTGAGCAATAATTA 3316
Db      |||||

```

Db 442 ACTTTTATAGAAATTAATTAATAAATTAATTTAACTTTAAATTCCTATTATTAATTT- 500
QY 3317 GGATATTTTGAATGGTTCATGAGATGCTTTTGTGCTGCTGCAATTTATGAGTTTATGA 3376
Db 501 TTATATATTTTAATAAATTAATTCATTTTATTTATTTATTTATTTATTTATTAATAATTA 560
QY 3377 TTTATGAATTAATGATGCTTTTTTTT 3401
Db 561 TTTAATTAATATTTATCATTTATTT 585

RESULT 42
US-10-204-708-71
; Sequence 71, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 71
; LENGTH: 8607
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-71

Query Match 2.2%; Score 78.8004; DB 1; Length 8607;
Best Local Similarity 45.4%; Pred. No. 25; Mismatches 377; Indels 4; Gaps 1;
Matches 317; Conservative 0;

QY 2790 GAACGTGACTGAAACCTTAGTGTATTTATTTACTCAGAAATAGTAATTTTCATATGTAAT 2849
Db 6156 GAGTGTATGTTATTTAGTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 6215
QY 2850 CAAATATTTTCATAGTTGGTTAGATTAATAGATTTTCAATTCGATTTTATCTTTG 2909
Db 6216 TAGAAATATAGATTTATTTGTTTATAGATTAATAGTTTATTTATTTGAAATTTTTCATGG 6275
QY 2910 ATTTTCTCTACTTATTTAATTTTGGATTTTAACTATTTTCTCAATGACTTGTATTTCT 2969
Db 6276 TATTTATGGATATTTAATTTATGATAAAGTGTTTTATTTTATTTTATTTTATTTTATTTT 6335
QY 2970 AATATTTACTTATCTATTTTACTTTTAAATGCACTTATTTTATTTGATTTTCTTAATAA 3029
Db 6336 AAAAGTAGTTTTTGTGGGTGTATTAATAAAGTAGTAGTATTTTGTGATTTTAA 6395
QY 3030 ATCCAGTCCCTGTTTTTTTTTAAAGACTTTAAATTAATTAATTTCTTTTATGTTTCTAC 3089
Db 6396 GTTTTTTATGTTATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 6455
QY 3090 CAGTCTTTTCCAGGCTACTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 3145
Db 6456 AAGGAATTTTGTGGTTTATTTAGTACGTTTTTTTTTTTTTAAATAATTTAATTTTATTTT 6515
QY 3146 GAATGCTAGTAACTCAATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 3205

Db 6516 AATTTTTAGTGTATTTGTTGTTGTAATTAATTTAGTGAATAAATTTTTTTTTTATGT 6575
QY 3206 TTTCTTTGATAACAGCTTTCAGTTCTATGGCTTTTAAATAAGTTTTTTTTTTTTTTTTT 3265
Db 6576 GTTGTGTGAGATAATAATTAATTTAGAAATTAATTTATTAATTTTGTAGAGAAATTTAT 6635
QY 3266 AAAGAAATGTCATTTCTTTGTGAAGTTTGTGCAATGCTTTGAGCAATAAATTTAGGATATTT 3325
Db 6636 TAGTTTTTTAAATAATTTTACGGTTAAGAGATTAATTTAAATAAGAGTAGTTTTTTAT 6695
QY 3326 TGAATGGTTCAAGATGATGCTTTTGTACTTGGCATTATTTAGAGTTTATGATTAAGAT 3385
Db 6696 GATTATTTTAAATGTTATTTTATTTTTCGTATGGAGTCGTATTTATTTTGTAT 6755
QY 3386 TATGATGCTTTTTTTTGGGCATAAAGGTCATGGCATATTTTGTGCTATATTTCTTA 3445
Db 6756 TTTTATTTGTTGATTTGTTGTTTTTTTGGTAGATTAATTTTATTAATAGTGTGTT 6815
QY 3446 AAATATAAATGCGTTTAAAGATTTTGTGCTATT 3483
Db 6816 AAATATTTGGGGGAGGTAGATACGGTGGTTTATTT 6853

RESULT 43
US-10-204-708-26
; Sequence 26, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 26
; LENGTH: 6583
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-26

Query Match 2.2%; Score 78.5002; DB 1; Length 6583;
Best Local Similarity 47.8%; Pred. No. 33; Mismatches 255; Conservative 0; Indels 3; Gaps 1;

QY 2868 TTGGTTAAGATAAATAAGATTTTCAATTTTATTTATTTTATTTTATTTTATTTTATTT 2927
Db 6052 TTGGTTTTTTTTTATTAATTAATTTATTTTGTGAGATTTATTTGATTTATTCGAGATTTA 6111
QY 2928 AATTTTGGATTTTAACTATTTCTCAATGACTTGTATTTCTAATATTTTACTTATTTCTAT 2987
Db 6112 TAGTTTGTTTTTTGTGCTATTTTAAATGAGAAATTTTATTAAGTAATTAATTAATAAT 6171
QY 2988 TTTACTTTTAAATGCACTTATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTT 3047
Db 6172 ATTTATTTTATTTATTTAGATTAATTTGAGTGGTTTTTAAAGATTTTATTTATTTATTT 6231
QY 3048 TAAAGAGACTTTAAATTAATTTATTTTCTTTAGTGTGTTTTACAGATTTCTTTCAGGCTACT 3107

Db 6232 TGAATATTTTGAATTTTAAGTGATGATGATTTTAAATTTTATTTTGTGTTTGT 6291
Qy 3108 TCTTTTGAATTTTGGTCTATCTTTTCTCAAGTTTGAATGGTACGTAACTCATTT 3167
Db 6292 TATTTTATTTAGTTATTTAGTTTTCGTTGTTATTTTGAATATGTTAGTATATTTT 6351
Qy 3168 ATCTTTATTTTGTAAATAGTCTTTTAAATTCATTTCTTTGTATACAGCTTCAGTT 3227
Db 6352 TTTTATAGGTTTGTATACGTTGTTTTTTTTTTTGAATGTTTTTTTTTAATAGATAGTT 6411
Qy 3228 CTATGGCTTTAATAAAGTTTTTTTTTTTTTTTTTTTTTAAAGAAATGCTATTTCTTGAA 3287
Db 6412 C---GTAGTTTGGTTTTTATTTTCGTTAGTAATATATATAGCTTTTTTTTAGTGAGAT 6468
Qy 3288 GTTTTGCAAGCTTTGAGCAATATTTAGGATATTTTGAATGGTTCAGGATAGCTT 3347
Db 6469 ATTTTATTAATATGTTATTTTAAATTTGTAATTTTATTTTATTTTGGGATATTTT 6528
Qy 3348 TTGTACTTGGCATTTATGAAAGTTTATGATTTATGATTTATGATGCTTTTTTT 3400
Db 6529 TTTTATTTTGAATTTATTTTATTTTATTTTATTTTATTTTATTTTGGATATATAT 6581

RESULT 44
US-08-998-416-288
; Sequence 288, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebschung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHYA GOSSYPIL
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwalis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 35,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 288:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 837 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:

ORGANISM: PAG124LRP
US-08-998-416-288
Query Match 2.2%; Score 78.3004; DB 1; Length 837;
Best Local Similarity 47.8%; Pred. No. 2.5e+02;
Matches 311; Conservative 0; Mismatches 332; Indels 7; Gaps 3;
Qy 2843 ATGTATTCAAAATATTTTTCATAATGTTGGTTAAGATAAAGATTTTCAAAATTCATTTT 2902
Db 22 ATTATTAAGATTATATATAAACTTTTATTATAATATTTAAGATTAAATTTATTTAAAC 81
Qy 2903 ATCTTTGATTTTCTCTACTTATTTAAATTTTGGGATTTTAACTATTCTTCAAGACTTG 2962
Db 82 TATTATTATCATTTATTAATAAATTAATTTATTTGATTATTAATCTTATTATAAATAT 141
Qy 2963 TATTTCTAATTTTACTTATTTTCTTTTAAATGCACTTATTTTATTTGATTTTTC 3022
Db 142 TATATAATTTACTTAACTCAATTTATTAATTTATTAATTAATAAATAAATATTAA 201
Qy 3023 T---AATAAAATCCAGTCTCTGTTT---TTTAAAGAGACTTTAAATATTAATTTCTC 3076
Db 202 TATGAATACATTTTAGTCTATGTTCAAAATTTTAAATTTAGTTTATTAATAATATTAGATA 261
Qy 3077 TTTAGTGTTTTACCAAGTTCTTTTCAGGCTACTCTTTTGATTTTATTTGGTCTATCTTTTC 3136
Db 262 TTATTATTTTCTTTTAAATAATTTATTAAGATTTATCAATTAATTAATTAATTTATTA 321
Qy 3137 TCAAGTTTGAATTTGGCTACGTAACTCATTTTATCTTTTATTTTGTAAATTTAGCTCTTTA 3196
Db 322 ATGTGTTTATAAATAATATATTTTATTTATTAATAAGATTTTAAATTTTAAATATTGTA 381
Qy 3197 AATTCATTTCTTTGATAACAGCTTCAGTTCTATGCTGTTTAAATAAGTTTTTTTTTTT 3256
Db 382 AATTATTTATTTTATTAATAATCTATTTTATAAATAATTAATGTTGATTATTTATTTA 441
Qy 3257 TTTTATTTTAAAGAAATGTCATTTCTTTGTGAAGTTTGTGCAATGCTTTGAGCAATAATTTA 3316
Db 442 ATCTTTTATAAGAAATATATATAAATAATTTTAACTTTAACTTTCTTATTTATTTAAT 501
Qy 3317 GGATATTTTGAATGGTTTCATGAGTATGCTTTTGTACTTTGGC-ATTATTTGAAGTTTATG 3375
Db 502 TTATATTTTAAATAAATATATTTTCAATTTTATTTTATTTTATTTTAAATAAATTAAT 561
Qy 3376 ATTTATGAATTTATGCTTTTGTGGGCAATAAGGCTATGGCATATTTTGTGGTC 3435
Db 562 ATTTAATTAATTTTATCAATTTTATTAATTAATAAATAATTAATAAAGATGAGTT 621
Qy 3436 TATATTTCTTAAATTAATAAATTTGGCTTTTAAAGATTTTGTGCTATTAA 3485
Db 622 AAAATACTTATAAAGGATCCGAACCTATATTATTGTTTATGAGACAAA 671
RESULT 45
US-10-204-708-39
; Sequence 39, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30

;; PRIOR APPLICATION NUMBER: DE 10043826.1
;; PRIOR FILING DATE: 2000-09-01
;; NUMBER OF SEQ ID NOS: 98
;; SEQ ID NO 39

;; LENGTH: 19513

;; TYPE: DNA

;; ORGANISM: Artificial Sequence

;; FEATURE:

;; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-204-708-39

```
Query Match          2.2%; Score 77.6002; DB 1; Length 19513;
Best Local Similarity 48.2%; Pred. No. 12;
Matches 327; Conservative 0; Mismatches 339; Indels 12; Gaps 4;

QY 2806 TTAGTAGTTTATATACCTCAGAAATAGTAAATTCATATGATGATTCATAATTAATTTTATATAA 2865
   |||||
DB 12679 TTATTAATTTTATTTTAAATTTAAGTTTAAATTTAGGATATATATATGTTATGTTAA 12738

QY 2866 TGTTCGTTAAGATATTAAGATTTTCAAAATGCAATTTTATCTTCATTTTCTCTACTTAT 2925
   |||||
DB 12739 TTTTATTATTTTAAAGATAATGTT---TGTGTTTTTTGTTTACATTTGATTTAGTGGT 12794

QY 2926 TTAATTTTGGGATTTTAACTATTTCTTCAATGACTGTTGATTTCTAATATTTACTTATCT 2985
   |||||
DB 12795 GTAGTTATGG--TTTATTTGAGTTTTTATTTTATGTTTAAAGTATTTTATTTTAG 12852

QY 2986 ATTTTACTTTAATGCACTTATTTTATTTATGATTTTCTAATAAATCCAGTCCTTGTGTTT 3045
   |||||
DB 12853 TTTTAAAGTATGTTGGGATTTAGGATGATTTATTTATGTTTAAAT--AAATTTTAAATTT 12910

QY 3046 TTTAAAGACCTTTAAATATTTATTTCTCTTTAGTGTGTTTACCAGTCTTTTCAGGCTA 3105
   |||||
DB 12911 TTGTAGAGATTTGATTTAAATTTTAAAT---TTTGTAGAAATGGAGTTTATATATGTTG 12966

QY 3106 CTCTTTTGAATTTTGGTCTATCTTTCTCAAGTTTTGAATGGCTACGTAACATCAT 3165
   |||||
DB 12967 TTTAGTTAGTTTATTTTAAATTTTGGTTTAAAGTGATTTATTTATTTAGTTTAAAGTG 13026

QY 3166 TTATCTTTATTTTGTAAATAGTCTTTTAAATTCATTTATCTTTGATTAACAGCTTCAG 3225
   |||||
DB 13027 TTGGATTTATAGACGTGAGTTTTTATGTTAGGTAATTTATTTTATGATATCGATAG 13086

QY 3226 TTCTATGGCTTTAATAAAGTTTTTTTTTTTTTTTTTTTTTAAAGAAATGTCATCTTTTGTG 3285
   |||||
DB 13087 TTTTAAAGATTATAGATTAGTTGTTGTTATTTTATTTTGGATTCGTTGTTTTTTTTTT 13146

QY 3286 AAGTTTTCACAATGCTTTGAGCAATAAATTTAGGATATTTTGAATGTTTCATGAGTATGC 3345
   |||||
DB 13147 TTGTTTTTGTGTTTGCAGATAGAGTTTGTGTTGTTATTTAGGTTGAGTGTAGTGGCGTAT 13206

QY 3346 TTTTGTACTTGGCATTTTATGAGTTTATGATTTTATGAATTTATGATCTTTTGTGGGC 3405
   |||||
DB 13207 CGTAGTTTATGTAATTTTGTGTTTTTATGTTTAAAGTATTTTGTGTTTTAGTTTTTTA 13266

QY 3406 ATAAGGCTATGTCATATTTTGTGCTATATTTCTTAAATATTAATTAATTCGCTTTAA 3465
   |||||
DB 13267 AGTAGTTGGTATTTAGGTAATTTATTTATTTAGTTTGGTTAAATTTTGTATTTTAGTAGAG 13326

QY 3466 AAAGTATTTGCTGCTATT 3483
   |||||
DB 13327 ATGGGGTTTTTATTTATGTT 13344
```

Search completed: August 24, 2004, 13:31:55
Job time : 614 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 13:34:23 ; Search time 279 Seconds
(without alignments)
3.941 Million cell updates/sec

Title: US-10-664-775-2

Perfect score: 3572
Sequence: 1 gtcaggaaggcgagctga.....gcaacacagcagaagctt 3572

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 0.5

Searched: 43 seqs, 153908 residues

Total number of hits satisfying chosen parameters: 86

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : rnpb2.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	3572	100.0	3572	1	US-10-664-775-2
2	1012.1	28.3	78056	1	Sequence 2, Appli
3	966.804	27.1	23651	1	Sequence 1, Appli
4	885.401	24.8	78056	1	Sequence 2, Appli
5	557.9	15.6	23651	1	Sequence 1, Appli
6	413.4	11.6	908	1	Sequence 1493, Ap
7	380.6	10.7	967	1	Sequence 11, Appl
8	372.6	10.4	419	1	Sequence 13832, A
9	364.7	10.2	451	1	Sequence 844, App
10	363.2	10.2	436	1	Sequence 6161, Ap
11	360.3	10.1	432	1	Sequence 5090, Ap
12	357.1	10.0	433	1	Sequence 4684, Ap
13	353.3	9.9	433	1	Sequence 5347, Ap
14	347.4	9.7	436	1	Sequence 7158, Ap
15	344.199	9.6	31412	1	Sequence 3, Appli
16	335.8	9.4	428	1	Sequence 8154, Ap
17	333	9.3	406	1	Sequence 1308, Ap
18	324	9.1	423	1	Sequence 8311, Ap
19	321.6	9.0	410	1	Sequence 13834, A
20	320.8	9.0	423	1	Sequence 5499, Ap
21	320.4	9.0	409	1	Sequence 14308, A
22	313.4	8.8	425	1	Sequence 9545, Ap
23	307.8	8.6	422	1	Sequence 3391, Ap
24	299.8	8.4	336	1	Sequence 11214, A
25	299.4	8.4	447	1	Sequence 11945, A
26	296	8.3	410	1	Sequence 4133, Ap
27	289.4	8.1	356	1	Sequence 3913, Ap
28	286	8.0	413	1	Sequence 13876, A
29	285.2	8.0	405	1	Sequence 9592, Ap
30	284.5	8.0	407	1	Sequence 13585, A
31	282.9	7.9	423	1	Sequence 13323, A
32	279.1	7.8	365	1	Sequence 13570, A
33	275.7	7.7	448	1	Sequence 10238, A

ALIGNMENTS

RESULT 1

US-10-664-775-2	7.7	430	1	US-09-960-352-6765	Sequence 6765, Ap
Sequence 2, Application US/10664775	7.6	436	1	US-09-960-352-9318	Sequence 9318, Ap
Publication No. US2004011576A1	7.6	396	1	US-09-960-352-11236	Sequence 11236, A
GENERAL INFORMATION:	7.6	449	1	US-09-960-352-7912	Sequence 7912, Ap
APPLICANT: Simesen, Ruth B	7.5	347	1	US-09-960-352-11744	Sequence 11744, A
APPLICANT: Pedersen, Anette A	7.5	383	1	US-09-960-352-7893	Sequence 7893, Ap
APPLICANT: Faisst, Steffen	7.5	432	1	US-09-960-352-2913	Sequence 2913, Ap
APPLICANT: Jensen, Jan J	7.4	347	1	US-09-876-143-32	Sequence 32, Appl
APPLICANT: Weiguny, Dietmar	7.2	423	1	US-09-960-352-14979	Sequence 14979, A
TITLE OF INVENTION: Method for Making Recombinant Proteins	7.2	417	1	US-09-960-352-149	Sequence 149, App
FILE REFERENCE: 6448.200-US	7.1	456	1	US-09-960-352-7209	Sequence 7209, Ap
CURRENT APPLICATION NUMBER: US/10/664,775	6.9	430	1	US-09-960-352-11719	Sequence 11719, A
PRIOR FILING DATE: 2003-09-17					
PRIOR APPLICATION NUMBER: Danish Application PA 2002 01384					
PRIOR FILING DATE: 2002-09-20					
PRIOR APPLICATION NUMBER: US 60/416,566					
PRIOR FILING DATE: 2002-10-07					
NUMBER OF SEQ ID NOS: 9					
SOFTWARE: PatentIn version 3.1					
SEQ ID NO 2					
LENGTH: 3572					
TYPE: DNA					
ORGANISM: Baby hamster kidney cell line					
US-10-664-775-2					
Query Match	100.0%	Score	3572;	DB 1;	Length 3572;
Best Local Similarity	100.0%;	Pred. No.	1.4e-13;		
Matches 3572;	Conservative	0;	Mismatches	0;	Gaps 0;
Qy	1	GTCCAGGAGGGCGGCGAGTGTGAGGAGGTACCTCCCTCGTCCAGGTAGGAGCAGTACG	60		
Db	1	GTCCAGGAGGGCGGCGAGTGTGAGGAGGTACCTCCCTCGTCCAGGTAGGAGCAGTACG	60		
Qy	61	TGGCGTTTGTCTGGAGCAGCGGTAAGAGATATCCCAAGGTAAGAGAAACCCCAAG	120		
Db	61	TGGCGTTTGTCTGGAGCAGCGGTAAGAGATATCCCAAGGTAAGAGAAACCCCAAG	120		
Qy	121	TAGATGTAGGTGTGTGAGAGGGGATCAGAGGGGAGACATCTGTAACCTCAATGAACTAAG	180		
Db	121	TAGATGTAGGTGTGTGAGAGGGGATCAGAGGGGAGACATCTGTAACCTCAATGAACTAAG	180		
Qy	181	GAAACTAGTCAATCTAATCACAAGTACAGGACCAAGCTGTCTAATCAATGAACTAAG	240		
Db	181	GAAACTAGTCAATCTAATCACAAGTACAGGACCAAGCTGTCTAATCAATGAACTAAG	240		
Qy	241	CCATGCCCTGGGGCAACCAAGATGGGAGGTCTGTGGAGAGATCTGACAGAAATGTG	300		
Db	241	CCATGCCCTGGGGCAACCAAGATGGGAGGTCTGTGGAGAGATCTGACAGAAATGTG	300		
Qy	301	GTCCACTGAGAGGAAGTCAACCACTTCAGTATCTTGTGCTTGTGAGAACCCCATGAAC	360		
Db	301	GTCCACTGAGAGGAAGTCAACCACTTCAGTATCTTGTGCTTGTGAGAACCCCATGAAC	360		
Qy	361	AGTATCAAAAGGCAAAATGATAGTACTGAAAGAGGAACTCCCAAGTCAAGTGTGCTG	420		

Db 361 AGTATGAAAGCCAAAATGATAGGATACCTGAAAGAGGAACTCCCCAGGTCAGTAGGTGCC 420
Qy 421 CCATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGAAATGAAGAGATGGAGCCA 480
Db 421 CCATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGAAATGAAGAGATGGAGCCA 480
Qy 481 AAGCAAAAAGAAATACCCAGCTGTGGATGTGATGTGATATAGCAAGGTCCGATGCTGT 540
Db 481 AAGCAAAAAGAAATACCCAGCTGTGGATGTGATGTGATATAGCAAGGTCCGATGCTGT 540
Qy 541 AAAGAGCAATATTCATAGGAACCTGGAAATGTCAGGTCCATGAATCAAGGCAAAATTCGAA 600
Db 541 AAAGAGCAATATTCATAGGAACCTGGAAATGTCAGGTCCATGAATCAAGGCAAAATTCGAA 600
Qy 601 GTGGTCAAAACAGAGATGGCAAGAGTGAATGTCAACATTCCTAGGAATCAGCGAACTAAAA 660
Db 601 GTGGTCAAAACAGAGATGGCAAGAGTGAATGTCAACATTCCTAGGAATCAGCGAACTAAAA 660
Qy 661 TGGACTGGAATGGGTGAATTTAACTCAGATCAACATTAATCTACTACTCGGGCAGGAA 720
Db 661 TGGACTGGAATGGGTGAATTTAACTCAGATCAACATTAATCTACTACTCGGGCAGGAA 720
Qy 721 TCCCTCAGAGAAATGGAGTAGCCATCATGTCTCAACAAAAGAGTCCGAAATGCAGTACTT 780
Db 721 TCCCTCAGAGAAATGGAGTAGCCATCATGTCTCAACAAAAGAGTCCGAAATGCAGTACTT 780
Qy 781 GGATGCACTCAAAACAGAGATGATCTCTGTTTGTTCAGGCAAAACCAATTCAT 840
Db 781 GGATGCACTCAAAACAGAGATGATCTCTGTTTGTTCAGGCAAAACCAATTCAT 840
Qy 841 ATCAGAGTAATCCAGTCTATGCCCAACCAAGTAACTGCTGAAGAGTGAAGTGAACGG 900
Db 841 ATCAGAGTAATCCAGTCTATGCCCAACCAAGTAACTGCTGAAGAGTGAAGTGAACGG 900
Qy 901 TCCTATGAGACCTTACAGACCTTTTGAAGTAAACACCCAAAAGAGTCTCTTCTCAT 960
Db 901 TCCTATGAGACCTTACAGACCTTTTGAAGTAAACACCCAAAAGAGTCTCTTCTCAT 960
Qy 961 ATAGGGAGTGAATGCAAAAGTAGGAGCAAGAAACACCTGAGTAAACAGGCAAAATTT 1020
Db 961 ATAGGGAGTGAATGCAAAAGTAGGAGCAAGAAACACCTGAGTAAACAGGCAAAATTT 1020
Qy 1021 GGCCTTGAATACGGAATGAAGCAGGCAAGAGTAAATAGAGTTTGGCCAAAGAAATGCA 1080
Db 1021 GGCCTTGAATACGGAATGAAGCAGGCAAGAGTAAATAGAGTTTGGCCAAAGAAATGCA 1080
Qy 1081 CTGGTATAGCAACACCTCTTCCACACACAGAGAGAGTCTACACATGGACATCA 1140
Db 1081 CTGGTATAGCAACACCTCTTCCACACACAGAGAGAGTCTACACATGGACATCA 1140
Qy 1141 CCAGATGGTCAACACCGAAATCAGATTAATTAATCTTTGAGCCCAAGAGATGGAGAGC 1200
Db 1141 CCAGATGGTCAACACCGAAATCAGATTAATTAATCTTTGAGCCCAAGAGATGGAGAGC 1200
Qy 1201 TCTATACAGTCAAGCAAAACAGACCCAGGAGCTTACTGGCTCAGATCATGAACTCCTT 1260
Db 1201 TCTATACAGTCAAGCAAAACAGACCCAGGAGCTTACTGGCTCAGATCATGAACTCCTT 1260
Qy 1261 ATTGCCAAATTCAGACTTAAATTAAGAAAGTAGGAAACCACTAGATCACTCAGGTAA 1320
Db 1261 ATTGCCAAATTCAGACTTAAATTAAGAAAGTAGGAAACCACTAGATCACTCAGGTAA 1320
Qy 1321 GACCTAAATCCATCCCTTATGATTAATACAGTGAAGTAGGAAATAGATTAAGGSCCTA 1380
Db 1321 GACCTAAATCCATCCCTTATGATTAATACAGTGAAGTAGGAAATAGATTAAGGSCCTA 1380
Qy 1381 GATCTGATGACAGAGTACCTTAATGAACTATGACAGAGGTTTCAATGATGTPACAGGAG 1440
Db 1381 GATCTGATGACAGAGTACCTTAATGAACTATGACAGAGGTTTCAATGATGTPACAGGAG 1440
Qy 1441 ACAGGATCGAGACCAATCCCATGGAAAGAAATGCAAAAAGCAAAATGGCTGTCTGG 1500
Db 1441 ACAGGATCGAGACCAATCCCATGGAAAGAAATGCAAAAAGCAAAATGGCTGTCTGG 1500

Qy 1501 GAGGCCCTTCAAAATAGCTGTGAAAAAGAGAGAGTGAAGCAAAAGGAAAAAGGAAAGA 1560
Db 1501 GAGGCCCTTCAAAATAGCTGTGAAAAAGAGAGAGTGAAGCAAAAGGAAAAAGGAAAGA 1560
Qy 1561 TAAAGCATCTGAATGCAGAGTTCMAAGAACTTCAGTTGTTCAAGCTGGTTTAGAAA 1620
Db 1561 TAAAGCATCTGAATGCAGAGTTCMAAGAACTTCAGTTGTTCAAGCTGGTTTAGAAA 1620
Qy 1621 AGTCAGAGCAACAGAGACCAAAATTCGCCAACTCTCTGTATCATGCAAAAAGCAAGAGA 1680
Db 1621 AGTCAGAGCAACAGAGACCAAAATTCGCCAACTCTCTGTATCATGCAAAAAGCAAGAGA 1680
Qy 1681 GTTCCAGAAAAACATCTATTTCTGCTTTATTCAGTATGCAAAAAGCTTTGACTGTGGGG 1740
Db 1681 GTTCCAGAAAAACATCTATTTCTGCTTTATTCAGTATGCAAAAAGCTTTGACTGTGGGG 1740
Qy 1741 TCACATTAACCTGTGAAAAATTTCTGAAAGGGATGGAAATACCAAGACCACTGACCTGACT 1800
Db 1741 TCACATTAACCTGTGAAAAATTTCTGAAAGGGATGGAAATACCAAGACCACTGACCTGACT 1800
Qy 1801 CTTGAAAAATTTGTATGCAAGGTCAAGAAAGCAACAGTTAGAACTGCACTGGAACAACAGA 1860
Db 1801 CTTGAAAAATTTGTATGCAAGGTCAAGAAAGCAACAGTTAGAACTGCACTGGAACAACAGA 1860
Qy 1861 CTGGTCCAGTAGGAAAAAGGATGATGCAAGGCTGATATTTGTCAACCGGCTTGTTTAA 1920
Db 1861 CTGGTCCAGTAGGAAAAAGGATGATGCAAGGCTGATATTTGTCAACCGGCTTGTTTAA 1920
Qy 1921 CTTCTATGCAGAGACATCATGAGAAACCTGGGCTGGAAAGAGCAACAAGCTGGAAATCAAG 1980
Db 1921 CTTCTATGCAGAGACATCATGAGAAACCTGGGCTGGAAAGAGCAACAAGCTGGAAATCAAG 1980
Qy 1981 ATTGCGGGAGAAATPAGCAATTAACCTCAGATATGCAAGATGATACCACTTATGCGAGAA 2040
Db 1981 ATTGCGGGAGAAATPAGCAATTAACCTCAGATATGCAAGATGATACCACTTATGCGAGAA 2040
Qy 2041 AGTGAAGAGGAACTAAAAAGCTCTTGTATGAAGGTGAAGAGGAGAGTGAAGAAAGTTGGC 2100
Db 2041 AGTGAAGAGGAACTAAAAAGCTCTTGTATGAAGGTGAAGAGGAGAGTGAAGAAAGTTGGC 2100
Qy 2101 TTAAAGCTCAACATTCAGAAAAACGAAGATCATGGCATCTGGTCCCATCACTTCATGGAA 2160
Db 2101 TTAAAGCTCAACATTCAGAAAAACGAAGATCATGGCATCTGGTCCCATCACTTCATGGAA 2160
Qy 2161 ATAGATGGGAAACAGTGGAAACAGTGTGAGAACTTATTTTGGGGGCTCCAAAATCA 2220
Db 2161 ATAGATGGGAAACAGTGGAAACAGTGTGAGAACTTATTTTGGGGGCTCCAAAATCA 2220
Qy 2221 CTGAGAGTGGTGACTGCGAGCCATGAAATTTAAAGACACTTACTCTCTGGAAAGAAAGTTA 2280
Db 2221 CTGAGAGTGGTGACTGCGAGCCATGAAATTTAAAGACACTTACTCTCTGGAAAGAAAGTTA 2280
Qy 2281 ACCACCTAGATAGCATATTCAGAAAGCAGACATTAACCTTGCCCAACAAAGCCCACTTA 2340
Db 2281 ACCACCTAGATAGCATATTCAGAAAGCAGACATTAACCTTGCCCAACAAAGCCCACTTA 2340
Qy 2341 GTCAAGGCTATGGTTTTCAGAGTGGTCAATGATGATGATGAGAGTGGAGTGTGAGAGTCA 2400
Db 2341 GTCAAGGCTATGGTTTTCAGAGTGGTCAATGATGATGATGAGAGTGGAGTGTGAGAGTCA 2400
Qy 2401 GCTGAGCACTGAAGAAATGATGCTTTTGAACCTGTTGGAGAGAGTCTTGGAGTCT 2460
Db 2401 GCTGAGCACTGAAGAAATGATGCTTTTGAACCTGTTGGAGAGAGTCTTGGAGTCT 2460
Qy 2461 CCTTGCATGCAAGAGATCCAAACAGTCCATTCGTAAGGAGATCAGCCCTGGGATTTCT 2520
Db 2461 CCTTGCATGCAAGAGATCCAAACAGTCCATTCGTAAGGAGATCAGCCCTGGGATTTCT 2520
Qy 2521 TTGAAAGGAATGATGCTAAAGCTGAAACTCCAGTACTTTTGGCCCACTGATCAGAAAGCT 2580
Db 2521 TTGAAAGGAATGATGCTAAAGCTGAAACTCCAGTACTTTTGGCCCACTGATCAGAAAGCT 2580


```
; SEQ ID NO 2
; LENGTH: 23651
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(23651)
; OTHER INFORMATION: n = A,T,C or G
US-09-884-586A-2

Query Match      15.6%; Score 557.9; DB 1; Length 23651;
Best Local Similarity 78.4%; Pred. No. 0.52;
Matches 820; Conservative 0; Mismatches 191; Indels 35; Gaps 13;

QY 1744 CAATAAAGCTGTGAAAAATCTGAAGGGATGGGAATACAGACCACCTGACCTGACTCTT 1803
DB 2456 CAAACAAAGCTGAAAAATCTTCAAGAGA--AGAATACAGACCACCTGACTCTTCT 2513
QY 1804 GAAAAATTTGTATGCAGGTCAGGAAGCAACAGTTAGAACTGGACATGGAACAACAGACTG 1863
DB 2514 GAGAAATCTGTTGTCTCA-GAAGCAACAGTTAGAACCAACAGATGGAACAACAGACTG 2572
QY 1864 GTTCAAGTAGGAAAGAGATGTCAAGGGCTGTATATTGTCACCCGGCTTTGTTAACTT 1923
DB 2573 GTTCCAAATCAGGAAGAGATGTCAAGGGCTGTATATCGTCACCCCTGATTATTAACTT 2632
QY 1924 CTATCGAGAG-ACATCATGAGAAAGCTGGGCTGGAAGAACACAGCTGAATCAAGAT 1982
DB 2633 ATATCGATAGTACATAATCAATTAAC-GAGATACAAGAT-ACACACACTTATGGCAGAAA 2750
QY 1983 TGCCGGGAAATAGCAATAACCTCAGATATGCAGATGATACACCCCTTATGGCAGAAAG 2042
DB 2693 TTCTGGGAAATATCAATTAAC-GAGATACAAGAT-ACACACACTTATGGCAGAAA 2750
QY 2043 TGAAGAGAACTAAAGACCTCTTGATGAAGTGAAGAGAGAGAGTGAAGAAAGTTGCTT 2102
DB 2751 CTAGAGAACTAAAGACCTCTTGATGAAGTGAAGAGAGAGAGTGAAGAAAGTTGCTT 2810
QY 2103 AAAGCTCAACATTCAGAAAGCAAGATCATGGCATCTGGTCCCATCACTTCATGGGAAT 2162
DB 2811 AAACCCCAACTTCAATCAAGA-----TCAATCACTTATGGGCAAT 2853
QY 2163 AGATGGGAAACAGTGGAAACAGTGCAGACTTTATTTTGGGGGCTCCAAATCACT 2222
DB 2854 AAATGGGAAACAAATGGAAACAGTGCAGACTTTATTTTCT--TGGGCTCCAAATCACT 2911
QY 2223 GCAGATGTGACTGCAGCCATGAATTAAGACACTTACTCTCTTGAAGAAAGTTA-- 2280
DB 2912 GCAGATGTGACTGCAGCCATG--ATTAAAGATGCTTGCTCTCTTGAAGAGAGAGTATT 2969
QY 2281 ACCAACCTTAGATACATATTGAAGAGAGAGACATTAACCTTGGCCAAACAAAGCCCCATCTA 2340
DB 2970 ACCAACTTAGAAGCATATTAAAGAGAGAGAGCGTTACTTTGCTGACTAAGTTCTGTCTA 3029
QY 2341 GTCAGGCTATGTTTTTCCAGTGGTCAATGATGATGATGAGAGTGGAGCTGGAAGAA 2400
DB 3030 GTCAAACCTATGTTTTTCCAGTGGTCAATGATGATGATGATGATGATGATGATGATGATG 3087
QY 2401 GCTGAGCACTGAAGATTTGATGCTTTTGAAGTGTGGTGTGGAGAGACCTTTGAGAGTC 2460
DB 3088 GCTGAGCAACCAAGATTTGATGCTTTTGAAGTGTGGTGTGGAGAGTCTTTGAGAGTC 3147
QY 2461 CTTTGA-CTGCAAGGAGATCCAAACAGTCCATTTCTGAAGAGATCAGCCCTGGGATTT 2519
DB 3148 CTTTGAACCTTGAAGGAGATCCAAACAGTCCATTTCTGAAGAGATCAGCCCTGAATTT 3207
QY 2520 TTTGGAAGGATGATGCTTAAGCTGAACTCCAGTACTTTGGCCACTGA--TCAGAAG 2577
DB 3208 ATTGGAAGGATGATGCTTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3267
QY 2578 GCTGACTCACTGAAAGACCTGTGCTGGAGGATTTGGGGGAGAGAGAGAGAGAGAGAGAG 2637
DB 3268 GCCAACTCACTAGAAAGAGCC-CATGTTGGCAAAATTTGAAGCCAGGAGAGAGAGAGAG 3326

QY 2638 CGACAGAGGATGAGATGGCTGGATGGCATCACTGACTCGATGACGCTGAGTCTGGGTGA 2697
DB 3327 TGACAGAGGATGAGATGGCTGGATGGCATCACTGACTCGATGACGCTGAGTCTGGATCA 3386
QY 2698 CTCTGAGTGGTGTGATGGACAGGAGGCTGCTCTCGGCGGATTCATGGGGTCAAAAAG 2757
DB 3387 TTCCGGGAGACAGCAAAAGGACAGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3446
QY 2758 AGTTGGACACGACTGAGCACTGAAC 2783
DB 3447 AGTCGGTCTCAAATGATGACTAAAC 3472

RESULT 6
US-09-876-143-1493/c
; Sequence 1493, Application US/09876143
; Publication No. US20040081958A1
; GENERAL INFORMATION:
; APPLICANT: Infogen Inc.
; APPLICANT: EILERTSEN, KENNETH J.
; APPLICANT: PEISTER-GENSKOW, MARTHA
; APPLICANT: CHILDS, LYNETTE
; APPLICANT: FORSYTHE, TODD
; APPLICANT: BISHOP, MICHAEL D.
; TITLE OF INVENTION: IDENTIFICATION AND USE OF MOLECULAR MARKERS INDICATING
; FILE OF INVENTION: CELLULAR REPROGRAMMING
; CURRENT APPLICATION NUMBER: US/09/876,143
; CURRENT FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/209,874
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 1744
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 1493
; LENGTH: 908
; TYPE: DNA
; ORGANISM: Bovine
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(908)
; OTHER INFORMATION: n is a, c, g, or t
US-09-876-143-1493

Query Match      11.6%; Score 413.4; DB 1; Length 908;
Best Local Similarity 85.6%; Pred. No. 18;
Matches 501; Conservative 0; Mismatches 68; Indels 16; Gaps 4;

QY 1663 CATGAAAAAGCAAGAGAGTCCAGAAAAACATCTATTCTGCTTTATTGACTATGCAAA 1722
DB 653 CATTGAAAAAGCAAGAGAGTCCAGAAAAACATCTATTCTGCTTTATTGACTATGCAAA 594
QY 1723 ACCCTT--GACTGTGGGGTCAATAACTGTGAAAAATCTGAAAGGATGGGAATA 1780
DB 593 AGGCTTTTGTGCTGATCAATAAATCTTATGAAAAATCTGAAAGAGATGGGATA 534
QY 1781 CCAGACCACTGACTGCTCTTGAATAATTTGATGAGCTGAGGAGCAAGCTAGTAGA 1840
DB 533 CCAGACCACTGATGCTCTTGAATAATCTTGAATAATCTTGAATAATCTTGAATAATCTTGA 474
QY 1841 ACTGACATGGAACAAACAGACTGGTTCCAGTGAAGAAAGAGTATGTCAGGCTGTATA 1900
DB 473 ACTGACATGGAACAAACAGACTGGTTCCAGTGAAGAAAGAGTATGTCAGGCTGTATA 414
QY 1901 TTGTCACCCGCTTGTATTAACTTCTATGTCAGAG-ACATCATGAGAAAGCTGGCTGAA 1959
DB 413 TTGTCACCCGCTTGTATTAACTTCTATGTCAGAGTACATCATGAGAAATCTGCTGGCTGGAG 354
QY 1960 GAAGCAAGCTGGAATCAAGATTCAGGAGGAGAAATAGCAATAACCTCAGATATGCAGAT 2019
DB 353 AAAGCAAGCTGATTCAGAGATTCAGGAGGAGAAATATCAATAACCTCAGATATGCAGAT 294
QY 2020 GATACCCTTTATGGCAGAAAGTGAAGAGAACTAAAAAGCCCTTTGATGAGGCTGAA 2079
```

```

Db      717  GACTGGATGCACAAGAGTTTGAGCAAGGTCGGGAGTTGGTAATGTCTAGGGAAGCCTCG 776
Qy      2731  CTTGCGCGGATTTCATGGGTGCACAAAGAGTTTGACACGACTGAGCAACTGAACCTGAACCTG 2790
Db      777  AGTGCTGCAGTCACAGGTCGCAAGAGTTCAGACACTACTGAGCGACTGAACCTGAACCTG 836
Qy      2791  AACTGTA 2797
Db      837  AACTGAA 843

RESULT 8
US-09-960-352-13832
; Sequence 13832, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Rao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathalagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21 (10298)/C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 13832
; LENGTH: 419
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 59-LIB3058-043-Q1-K1-G12
US-09-960-352-13832

Query Match          10.4%; Score 372.6; DB 1; Length 419;
Best Local Similarity 93.1%; Pred. No. 44;
Matches 390; Conservative

Qy      2352  GGTGTTTCCAGTGGTCATGATGATGTGAGAGTTGGAGTGTGAAGAAAGCTGAGCACTG 2411
Db      1      GGTGTTTCTGTGGTCATGATGATGATGTTAGAGTTGGACTGTCAAGAAGGCTGAGTGCGG 60
Qy      2412  AAGAATTGATGCTTTTGAAGTGTGGTGTGGAGAGACTCTTGAGAGTCCCTTGGACTGC 2471
Db      61  AAGAATTGATGCTTTTGAAGTGTGGTGTGGAGAGACTCTTGAGAGTCCCTTGGACTGC 120
Qy      2472  AAGGAGATCCAAACAGTCCATCTTGAAGGAGATCAGCCCTGGGATTTCTTTGGAAGGAAT 2531
Db      121  AAGGAGATCCAAACAGTCCATCTTGAAGGAGATCAGCCCTGGGATTTCTTTGGAAGGAAT 180
Qy      2532  GATGCTAAAGCTGAACCTCCAGTACTTTGCCACCTGATCAGAGAGCTGACTCCTGGA 2591
Db      181  GATGCTGAAGCTGAATTCAGTACTTTGCCACCTCATCGAAGAGTTGACTTATGGA 240
Qy      2592  AAAGACCTTGATCTCGGAGGGATTGGGGCAGGAGGAAGGGGACGACAGAGGATGAG 2651
Db      241  AAAGACTCTGATCTCGGAGGGATTCGGGGCAGGAGGAAGGGGACGACAGAGGATGAG 300
Qy      2652  ATGGCTGGATGGCATCACTGACTGATGACCTGAGTCTGGGTGAACCTCCTGGAGTTGTT 2711
Db      301  ATGGCTGGATGGCATCACCGACCTGATAGAGTGAAGTCTGAGTGAACCTCGGGAGTTGTT 360
Qy      2712  GATGGACAGGAGGCTGCTGCTCGGCCGATTCATGGGGTCAAAAGAGTTGACACGAC 2770
Db      361  GATGGAAAGGAGGCGCTGCGCTGCTCGGATTCATGGGGTCGCAAAAAGTTGGACACGAC 419

RESULT 9
US-09-960-352-844
; Sequence 844, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.

```

26 GATCTTCAGTTGTTCAAGCTGGTTTGTAGAAAAGGCAGAGGAACCAATGCC 85

26 GATCTCCAGTTGTTCAAGCTGGTTTTAGAAAAGGCAGAGGAACCAAGATCAAATGGC 85

db 26 GATCTTCCAGTTGTTTCAAGCTGGTTTTAGAAAAGGCAGAGAACCCAGAGATCAAAATTGCC 85

QY 1649 AACATCCTCTGTATCATGGAAGCAAGAGAGTTCCAGAAAAACATCTATTCTCTCTTT 1708
| | | | |
Db 86 AACATCCGCTGAATCATCGAAGAAAGCAAGAGAGTTCCAGAAAAACATCTATTCTCTCTTT 145
| | | | |
QY 1709 ATTGACTATGAAAGCCTTTGACTGTGGGGGTGCACATTAACCTGTGGAAATTCAGAA 1768
| | | | |
Db 146 ATTGACTATCCAGGCGCTTGTGCTGTGGATCACATTAACCTGTGGAAATTCAGAA 205
| | | | |
QY 1769 GGGATGGGAATACCAGACCACTGACCTGACTCTTGAATAATTTGATGAGGTCAAGAA 1828
| | | | |
Db 206 AAGATGGGAATACCAGACCACTGACCTGCTCTTGAGAAATCTGTATGAGCCAGGAA 265
| | | | |
QY 1829 GCAACAGTTAGACTGACATGGAACCAACAGACTGGTTCCAAAGTAGAAGAGGATGATGT 1888
| | | | |
Db 266 GCAACAGTTAGACTGACATGGAACCAACAGACTGGTTCCAAAGTAGAAGAGGATGATGT 325
| | | | |
QY 1889 CAAGGCTGTATTGTCAACCGGCTTGTAACTTCTATCATGACAG-ACATCATGAGAAAC 1947
| | | | |
Db 326 CAAGGCTGTATTGTCAACCGCTTGTAACTTCTATCATGACAGTACATATGAGAAAC 385
| | | | |
QY 1948 GCTGGCTGGAAGACACAAAGCTGGAATCAAGATTGCCGGGAGAA 1994
| | | | |
Db 386 TCTGGCTGGAAGACACAAAGCTGGAATCAAGATTGCCGGGAGAA 432
| | | | |
RESULT 12
US-09-960-352-4684
; Sequence 4684, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wyatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 4684
; LENGTH: 433
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 20-LIB34-039-Q1-E1-E7
US-09-960-352-4684
Query Match 10.0%; Score 357.1; DB 1; Length 433;
Best Local Similarity 90.5%; Pred. No. 50;
Matches 401; Conservative 0; Mismatches 29; Indels 13; Gaps 2;
QY 1943 GAAACGCTGGCTGGAAGAACACACAGCTGGAATCAAGATTGCCGGAGAAATAGCAATA 2002
| | | | |
Db 1 GAAACGCTGGCTGGAGGAAGCAACAGCTGGAATCAAGATTGCCGGAGAAATATCAATA 60
| | | | |
QY 2003 ACCTCAGATATGCAGATGATACCACTTATGCGAGAAAGTGAAGAGAACTAAAAAGCC 2062
| | | | |
Db 61 ACCTCAGATATGCAGATGATACCTTATGCGAGAAAGTGAAGAGAACTAAAGAGCC 120
| | | | |
QY 2063 TCTTGATGAAGGTGAAAGAGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAAA 2122
| | | | |
Db 121 TCTTGATGAAGGTGAAAGAGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAAA 180
| | | | |
QY 2123 CGAAGATCATGGCATCTGGTCCCATCTTCATCGGAAATAGATGGGGAACATGTGAAA 2182
| | | | |
Db 181 CTAAGATCATGGCATCTGGTCCCATCTTCATCGGAAATAGATGGGGAACATGTG--- 237
| | | | |
QY 2183 CAGTGTACAGACTTTATTTTTGGGGGCTCCAAATCACTCAGATGGTGACTCAACCCA 2242
| | | | |
Db 238 -----CTGACTTTATTTTGGGGGCTCCAAATCACTCAGATGGTGACTCAACCCA 289
| | | | |
QY 2243 TGAATTTAAAGACATCTACTCTCTGGNAGAAAAGTTA-ACCACTAGATAGCATATT 2300
| | | | |

Db 290 TGAATTTAAAGACACTTACTCTTTGGAAGAAAGTTATGACCAACATAGACAGCATATT 349
| | | | |
QY 2301 GAAAAGCAGAGACATTACTCTTGGCCAAACAAAGCCCATCTAGTCAAGGCTATGTTTTTCC 2360
| | | | |
Db 350 AANTAGCAGACATCACTTTGTCAACAAAGGTCCATCTAGTCAAGGCTATGTTTTTTC 409
| | | | |
QY 2361 AGTGTCTATGATGATGTGAGA 2383
| | | | |
Db 410 AGTGTCTATGATGATGTGAGA 432
| | | | |
RESULT 13
US-09-960-352-5347
; Sequence 5347, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wyatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 5347
; LENGTH: 433
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 23-LIB3058-048-Q1-K1-F3
US-09-960-352-5347
Query Match 9.9%; Score 353.3; DB 1; Length 433;
Best Local Similarity 90.0%; Pred. No. 52;
Matches 389; Conservative 0; Mismatches 42; Indels 1; Gaps 1;
QY 1701 TCTGCTTTATTGACTATGCAAAAGCCTTTGACTGTGGGGTGCACATAAAGCTGTGAAAA 1760
| | | | |
Db 2 TCTGCTTTATTGACTATGCAAAAGCTTTTGAAGTCTTGTGATCACAATAAAGCTGTGAAAA 61
| | | | |
QY 1761 TTCTGAAAGGATGGGAATACCAGACCACTGACCTGCTTTGAAAAATTTGTATGCAG 1820
| | | | |
Db 62 TTCTGAAAGATGGGAATACCCTGCCACCTGCTTTGAGAAAACTTATGTCAG 121
| | | | |
QY 1821 GTCAGGAAGCAACACTGTAGAACTGGACATGGAAACACAGACTGTTCCAAAGTAGAAAAAG 1880
| | | | |
Db 122 GTCAGGAAGCAACACTGTAGAACTGGACATGGAAACACAGACTGTTCCAAATAGAAAAAG 181
| | | | |
QY 1881 GAGTATGTCAAGGCTGTATATTGTCAACCGGCTTTGTTAACTTTCTATGCAGAG-ACATCA 1939
| | | | |
Db 182 GAGTATGTCAAGGCTGTATATTGTCAACCGCTTTTAACTTTATGTACAGAGTACATCA 241
| | | | |
QY 1940 TGAGAAACCTGGCTGGGAAGAGCAACAGCTGGAAATCAAGATTGCCGGGAGAAATAGCA 1999
| | | | |
Db 242 TGAGAAATCTGGACTGGATGAAGCCCAAGCTGGAGTCAAAAATTCGCGGAGAGATATCA 301
| | | | |
QY 2000 ATAACTCTCAGATATGCAGATGATACCACCTTTATGCGAGAAAGTGAAGAGGAACATAAAA 2059
| | | | |
Db 302 ATAACTCTCAGATATGCAGATGATACCACCTTTATGCGAGAAATGAAGAGGAATTAATA 361
| | | | |
QY 2060 GCTCTTTGATGAAGGTGAAAGAGGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGA 2119
| | | | |
Db 362 GCCTCTTGATGAAGTGAAGAGAGAGTGAAGAAAGTTGCTTTAAAGCTTCACATTCAGA 421
| | | | |
QY 2120 AAACGAGATCA 2131
| | | | |
Db 422 AAACGAGATCA 433
| | | | |
RESULT 14
US-09-960-352-7158
; Sequence 7158, Application US/09960352

```

; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960.352
; NUMBER OF SEQ ID NOS: 24
; SEQ ID NO 7158
; LENGTH: 436
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 31-LIB2809-013-Q1-E1-H3
; US-09-960-352-7158

Query Match          9.7%; Score 347.4; DB 1; Length 436;
Best Local Similarity 92.8%; Pred. No. 54;
Matches 385; Conservative 0; Mismatches 26; Indels 4; Gaps 2;

QY 1931 GAGACATCATGAGAAACCTGGCTGGAGGAAAGTAAAGCTGGAATCAAGATTGCCAGGA 83
DB 24 GATACATCATGAAAAACCTGGCTGGAGGAAAGTAAAGCTGGAATCAAGATTGCCAGGA 83
QY 1991 GAAATAGCAATAAATCTCAGATATCAGATGATACACCTTATGCGCAGAAAGTGAAGAGG 2050
DB 84 GAAATATCAATAGCTTGATATGAGATGACACCTTATGCGCAGAAAGTGAAGAGG 143
QY 2051 AACTAAAAAGCTCTTGATGAAGGTGAAGAGGAGAGTGAAGAAAGTTGGCTTAAAGCTCA 2110
DB 144 AACTAAAGAGCTCTTGATGAAGGTGAAGAGGAGAGTGAAGAAAGTTGGCTTAAAGCTCA 203
QY 2111 ACATTCAGAAACGAGATCATGCTGCTGCTCCATCCTTATGCGCAGAAAGTGAAGAGG 2170
DB 204 ACATTTAGAAACGAGATCATGCTGCTGCTCCATCCTTATGCGCAGAAAGTGAAGAGG 263
QY 2171 AAACAGTGGAAACAGTGTGACAGCTTTATTTTGGGGGGTCCAAATCATCTGCAGATGG 2230
DB 264 AAACAGTGGAAACAGTGTGACAGCTTTA--TTTGGGGGGCTCCAAATCATCTGCAGATGG 321
QY 2231 TGACTGAGCCATGAATTAAGACACATCTACTCTTGGAGAAAGTTA--ACCAACCT 2288
DB 322 TGATTGAGCCATGAATTAAGACACATCTACTCTTGGAGAAAGTTAATGACCAACCT 381
QY 2289 AGATAGCATATTGAAAGCAGAGACATTAACCTTCCCAACAAAGCCCATCTAGTC 2343
DB 382 AGATAGCATATTGAAAGCAGAGACATTAACCTTCCCAACAAAGTCCATCTAGTC 436

RESULT 15
US-10-109-551-3
; Sequence 3, Application US/10109551
; Publication No. US20020194635A1
; GENERAL INFORMATION:
; APPLICANT: DUNNE, PATRICK W.
; APPLICANT: PIEDRAHITA, JORGE
; TITLE OF INVENTION: TRANSGENIC ANIMALS RESISTANT TO TRANSMISSIBLE
; FILE REFERENCE: SPONGIFORM ENCEPHALOPATHIES
; FILE REFERENCE: TAMK:207US
; CURRENT APPLICATION NUMBER: US/10/109.551
; CURRENT FILING DATE: 2002-03-28
; PRIOR FILING DATE: 60/280,549
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 31412
; TYPE: DNA
; ORGANISM: Ovis aries
; US-10-109-551-3

```

```

Query Match          9.6%; Score 344.199; DB 1; Length 31412;
Best Local Similarity 69.7%; Pred. No. 3.1;
Matches 492; Conservative 0; Mismatches 208; Indels 6; Gaps 2;

QY 2410 TGAAGAAATTGATGCTTTTCAACTGTGTTGGAGAAGACTCTTGAGAGTCCCTTGGACT 2469
DB 17733 TGNATACAGTAGGTTTGTGAATGTGTTATGGAGAGACTCTTGAGAGTCCCTTGGACT 17792
QY 2470 GCAAGAGATCCAAACAGTCCATTCTGAAGAGATCAGCCCTGGGATTTCTTTGGAAGA 2529
DB 17793 GCAAGGAGATCCAAACCGGTCCATTCTAAAGGAGATCAGCCCTGGAAATTTCTTTGGAAGAA 17852
QY 2530 ATGATGCTAAAGCTGAACTCCAGTACTTTGGCCACCTGATCAGAGAGCTGACTCACTG 2589
DB 17853 ATGATGCTAAAGCTGAAATCCAGTACTTTGGCCACCTCATGCAAGAGTTGACTCATG 17912
QY 2590 GAAAGACCCCTGATGCTGGAGGGGATTGGGGGCGAGAGAGAGGGGACGACAGAGGATG 2649
DB 17913 GAAAGACTCTGATGCTGGAGGGGATTGGGGGCGAGAGAGAGGGGACGACAGAGGATG 17972
QY 2650 AGATGGCTGGATGGCATCACTGCTGATGAGCTGAGTCTGGTGAACCTCTGAGTTG 2709
DB 17973 AGATGGCTGGATGGCATCACTGCTGATGAGCTGAGTCTGGTGAACCTCTGAGTTG 18032
QY 2710 GTGATGGACAGGAGGCGCTGCTCTGGCGGATTCTATGGGGTCACAAAGAGTTGGACACGA 2769
DB 18033 TTGATGGACAGGAGGCGCTGCGGTGCTGCAATTCATGGGTACAAAGAGTGGACACGA 18092
QY 2770 CTGAGCACTGAACCTGAACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2829
DB 18093 CTGAGTGAACCTGAACCTGAACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 18152
QY 2830 ATAGTAATTTTCATATGTTTCAAAATTTTCAATAATTTTCAATAATTTTCAATAATTTT 2889
DB 18153 TTTGTTATTTTGGACATTTCTTCACTCTT--ATTGAATTCCTCAAAAAAATCCAGC 18211
QY 2890 CAAATGATTTTATCTTTTGTATTTTCTCTACTTATTTTAAATTTTGGGATTTTAACTATT 2949
DB 18212 TAGAATTTTGTATTTTCAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 18271
QY 2950 CTTCAATGATGTTATTTCTAATTTTATTTATTTATTTATTTATTTATTTATTTATTTAT 3004
DB 18272 TTATTAATGATGTTATTTCTAATTTTATTTATTTATTTATTTATTTATTTATTTATTTAT 18331
QY 3005 TATTTTATTTATTTTCTAATAAAATCCAGTCTCTTTTATTTTATTTTATTTTATTTTATTT 3064
DB 18332 CTTTAGTAGTGTTCAGTTTCTTTTCCCGAGTCTTACITTTCTTTTGTAGTTAAGTCAATCC 18391
QY 3065 TATTTAATTTCTTTTAGTGTTTTACCAGTCTTTTCCAGCTACTTCT 3110
DB 18392 TAGATACCTTTACAGTATGATAGTGAAGTTGTATCTGACTCTTTGT 18437

RESULT 16
US-09-960-352-8154
; Sequence 8154, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960.352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 8154
; LENGTH: 428
; TYPE: DNA
; ORGANISM: Bos taurus

```


NAME/KEY: unsure
LOCATION: (31)
OTHER INFORMATION: unsure at all n locations
OTHER INFORMATION: Clone ID: 61-LIB3058-039-Q1-K1-H2
US-09-960-352-14308

Query Match 9.0%; Score 320.4; DB 1; Length 409;
Best Local Similarity 89.3%; Pred. No. 74;
Matches 366; Conservative 0; Mismatches 42; Indels 2; Gaps 2;

Qy 1748 AAACGTGGGAAATTCGAAAGGGATGGGAATACCAAGACACCTGACCTGACCTTTGAAA 1807
Db AAACGTGGGAAATTCGAAAGGGATGGGAATACCAAGACACCTGACCTTTGAAA 1807
Qy 1808 AATTTGTATCAGGTGAGGAAGCAACAGTTAGAACTGGACATGGAACACAGACTGGTTC 1867
Db AATTTGTATCAGGTGAGGAAGCAACAGTTAGAACTGGACATGGAACACAGACTGGTTC 1867
Qy 349 AATCTGTATCAGGTGAGGAAGCAACAGTTAGAACTGGACATGGAACACAGACTGGTTC 290
Db AATCTGTATCAGGTGAGGAAGCAACAGTTAGAACTGGACATGGAACACAGACTGGTTC 290
Qy 1868 CAAGTAGGAAGAGATGCTCAAGGCTGTATTTGTCACCGGCTTTGTTAACTTCTAT 1927
Db CAAGTAGGAAGAGATGCTCAAGGCTGTATTTGTCACCGGCTTTGTTAACTTCTAT 1927
Qy 1928 GCAGAG-ACATCATGAAACGCTGGGCTGGAAGAACACAACTGGAATCAAGATTGCC 1986
Db GCAGAGTCCATGATGAAACGCTGGGCTGGAAGAACACAACTGGAATCAAGATTGCC 1986
Qy 1987 GGGAGAAATAGCAATACCTCAGATATGAGATGATACCACTTATGGCAGAACTGAA 2046
Db GGGAGAAATAGCAATACCTCAGATATGAGATGATACCACTTATGGCAGAACTGAA 2046
Qy 2047 GAGGAATCTAAAGACCTCTTGATGAGGTGAAAGAGAGAGTGAAGAAAGTTGGCTTAAAG 2106
Db GAGGAATCTAAAGACCTCTTGATGAGGTGAAAGAGAGAGTGAAGAAAGTTGGCTTAAAG 2106
Qy 2107 CTCAACATTCAGAAACAGAGATCATGGATCTGGTCCCATCTACTTCATG 2156
Db CTTCACCTTCAGAAATG-NGATCATGGATCGGTCCTCCCTCCGTTCCATG 1

RESULT 22

US-09-960-352-9545
Sequence 9545, Application US/09960352
Patent No. US20020137139A1

GENERAL INFORMATION:

APPLICANT: Warren, Wesley C.
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960.352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112

SEQ ID NO 9545

LENGTH: 425

TYPE: DNA

ORGANISM: Bos taurus

OTHER INFORMATION: Clone ID: 41-LIB3058-043-Q1-K1-C10

US-09-960-352-9545

Query Match 8.8%; Score 313.4; DB 1; Length 425;
Best Local Similarity 87.4%; Pred. No. 77;
Matches 375; Conservative 0; Mismatches 46; Indels 8; Gaps 3;

Qy 2073 GGTGAAGAGGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAACGAAAGATCAT 2132
Db GGTGAAGAGGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAACGAAAGATCAT 2132
Qy 2133 GGCATCTGCTCCCATCTCATTTGGAATATAGATGGGAAACAGTGGAAACAGTGTGAGA 2192
Db GGCATCTGATCCCATCTCATTTTTCATGCAATATAGATGGGATCAATGGAACAGTGTGAGA 2192

Qy 2193 CTTTATTTTTTGGGGGCTCCAAATCACTGAGATGGTACTGACGACCATGAATTTAA 2252
Db CTTTATGTTT--TGGGCTCCAAATCACTGAGATGGTGAATGAGCCATGAATTTAA 180
Qy 2253 AGACACTTACTCCTTTGGAAGAAAAGTTA--ACCAACCTTAGATAGCATATGAAAAGCAGA 2310
Db TGACTCTTCTCTTTGGAAGAAAAGCTATGACCGAGCTAGACAGCATATTTAAAAAGCAGA 240
Qy 2311 GACATTACTTGCACAAAGCCCATCTAGTCAAGGCTATGGTTTTCAGTGGTCTATG 2370
Db AACATTACTTTACCAACAAGGTCCATCTATTTCAAAGCTA-----TTTTCTAGTAGTCTATG 296
Qy 2371 TATGATCTGAGAGTTGCACTGTGAAGAAAGCTGAGCACTGAAGAAATGATGCTTTTGAA 2430
Db TATGTTCTGAGAGTTGCACTATTAAGAAAGCTGAGTCTCTGAAGAACTGATGTTTGA 356
Qy 2431 CTGCTGGTTGGAAGAACTCTTGAAGTCCCTTGGACTCGAAGAGATCCAAACAGTCC 2490
Db CTGCGGTGTTGGAAGAACTCTGGAAGTCCCTTGAAGTCCCTTGAAGAGATCCAAACAGTCA 416
Qy 2491 ATTCTGAAG 2499
Db ATCTTAAG 425

RESULT 23

US-09-960-352-3391
Sequence 3391, Application US/09960352
Patent No. US20020137139A1

GENERAL INFORMATION:

APPLICANT: Warren, Wesley C.
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960.352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112

SEQ ID NO 3391

LENGTH: 422

TYPE: DNA

ORGANISM: Bos taurus

OTHER INFORMATION: Clone ID: 15-LIB3058-036-Q1-K1-D11

US-09-960-352-3391

Query Match 8.6%; Score 307.8; DB 1; Length 422;
Best Local Similarity 86.0%; Pred. No. 81;
Matches 363; Conservative 0; Mismatches 57; Indels 2; Gaps 2;

Qy 1714 CTATGCAAAAGCCTTTGACTGTGGGGTCAATAACTGTGGAATAATCTGAAAGGAT 1773
Db CTGCGTGAAGAACTTTGACTGTGTAGATCAACAATCTGTGGAATAATCTGAAAGAT 60
Qy 1774 GGAATACCAAGACCACTGACTCTTGAATAATTTGATGAGGTGAGGAGCAAC 1833
Db GGAATACCAAGACCACTGACTCTTGAATAATTTGATGAGGTGAGGAGCAAC 120
Qy 1834 AGTTAGAACTGGACATGGAACCAACAGACTGGTTCCTCAAGTGAAGAAAGAGTATGTCAAG 1893
Db AGTTAGAACTGGACATGGAACCAACAGACTGGTTCCTCAAGTGAAGAAAGAGTATGTCAAG 180
Qy 1894 CTGTATTTGTACCCGGCTTGTTHAATCTTATGTCAGAG-ACATCATGAGAACTGG 1952
Db CAATATTTGTCACTCCCTGCTTATTTAACTTATATGAGAGTACATCATGAGAAATCTG 240
Qy 1953 GCTGGAAGAGCAACAGCTGGAATCAAGATTGC-CCGGAAGAAATAGCAATTAACCTCAGAT 2011
Db GCTGGAATGAACCAACAGCTGGAATCAAGATTGCTGGGGAAGATATCAATTAACCTCAGAT 300
Qy 2012 ATGCAGATGATACCACTTATGGCAGAAAGTGAAGAGAACTTAAAGACCTTTCATGA 2071

Query Match 8.3%; Score 296; DB 1; Length 410;
Best Local Similarity 86.9%; Pred. No. 92;
Matches 358; Conservative 0; Mismatches 50; Indels 4; Gaps 3;

QY 593 AATTGGAAGTGGTCAACAAGAGATGCGAAGAGTGAATGCAACATTTCTAGGAATCAGCG 652
DB 410 AATTGGAAGGTTCAACACGAGATGCGAAGATGAACATCAACATTTTAGGATCACTG 351
QY 653 AACTAAATGAGCTGGATGGGTGAATTTAACTCAGATGCCATT-ATATCTACTACTGC 711
DB 350 AAATAAATGAATGGAATGGGTGAATATAACGCGATGACCAATTAATCTACTACTGT 291
QY 712 GGGCAGAAATCCCTCAGAGAAATGGAGTACCATCATGCTCAACAAAAGAGTCCGAAT 771
DB 290 GGGCAAGATCCCTTAGAAGAAATGGAGTAGCCATCATATCAACGTAGAGTCTGACAT 231
QY 772 GCACTACTGGATGCACTCTCAAAAACGACGAGATGATCTCTGTTTTCAGAGGCAAA 831
DB 230 GCA--ACTTGGATGCAATCTCAAAAACGACGAGATGATTTCT-ATTCATTTCCAAAGGCAA 174
QY 832 CCATTCATATCAGTAATCCAGTCTATGCCCAACCAAGTATGCTGAAGAAAGCTGAA 891
DB 173 CCATTCATATCAGTAATCCAGTCTATGCCCAACCAAGTATGCTGAAGAAAGCTGAT 114
QY 892 GTTCAACGGTCCCTATGAAGACCTCAAGACCTTTTGAAGTAAACCCCAAAAAGATGTC 951
DB 113 GGTGAATGGTCTATGAAGACCTCAAGACCTTTTGAAGTAAACCCCAAAAATATGTC 54
QY 952 CTTCTCATATAGGGGACTGGAATGCAAAAGTAGGAAGCAAGAACACCTG 1003
DB 53 CTTTTCATATAGTGGACTGGATGCAAAAGTAGGAATGCAAGAGATATATG 2

RESULT 27
US-09-960-352-3913/C
; Sequence 3913, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960.352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 3913
; LENGTH: 356
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 17-LIB34-010-Q1-E2-E1
US-09-960-352-3913

Query Match 8.1%; Score 289.4; DB 1; Length 356;
Best Local Similarity 91.6%; Pred. No. 11e+02;
Matches 327; Conservative 0; Mismatches 26; Indels 4; Gaps 2;

QY 2183 CAGTGTGACACTTTATTTTGGGGGGCTCCAAATCACTGCAGATGGTGAAGCCCA 2242
DB 356 CAGTGTGACACTTTA--TTTGGGGGGCTTCAATCACTGCAGATGGTGAAGCCCA 299
QY 2243 TGAATTTAAAGACACTTACTCTTTGGGAAGAAAGT--TAACCAACCTAGATGATATT 2300
DB 298 TGAATTTAAAGACACTTACTCTTTGGGAAGAAAGT--TAACCAACCTAGATGATATT 239
QY 2301 GAAAGCAGAGACATTTACTCTTCCCAAGACCCCTAGTCAAGGCTATGTTTTC 2360
DB 238 CAAAGCAGAGACATTTACTTTTCCCAAGAGAGTCTGTCTTAATCAAGGCTATGTTTTC 179
QY 2361 ACTGCTCATGTATGTATGTAGATGTGAGATGTGAGATGTGGAAGAAAGCTGACACTCAAGAATTGA 2420

DB 178 AGTGTCTATGTATGTATGTAGATGTGAGATGTGAGAAAGCTGAGTCTCAAGAATTGA 119
QY 2421 TCGTTTTCAACTGTGGTGTGGAGAGACTCTTGGAGTCTCCTTGGACTGCAAGGAGATC 2480
DB 118 TCGTTTTCATCTGTGGTGTGGAGAGACTCTTGGAGTCTCCTTGGATGCAAGGAGATC 59
QY 2481 CAACCAAGTCCATTTCTGAAGAGATCAGCCCTGGGATTTCTTTGGAAGGAATGATCT 2537
DB 58 CAACCAAGTCCATTTCTGAAGAGATCAGCCCTGGGATTTCTTTGGAAGGAATGATCT 2

RESULT 28
US-09-960-352-13876
; Sequence 13876, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960.352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 13876
; LENGTH: 413
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 59-LIB34-027-Q1-E1-E4
US-09-960-352-13876

Query Match 8.0%; Score 286; DB 1; Length 413;
Best Local Similarity 87.6%; Pred. No. 1e+02;
Matches 345; Conservative 0; Mismatches 45; Indels 4; Gaps 3;

QY 1889 CAAGGCTGTATTTGTACCCCGCTTGTAACTTCTTATGCAGAG-ACATCATGAGAAAC 1947
DB 1 CAAGGCTGTATTTGTCAACTCTCTTATTTAACTTATATTCAGAGTACATCATGCAAT 60
QY 1948 GCTGGGCTCGAAGAACACAAAGCTGGAATCAAGATTCGCGGAGAAATAGCAATACCTC 2007
DB 61 CTTGGGCTCGATGAAGACCAACCTGGAATCAAGATTCGCGGAGAAATATCAATACCTC 120
QY 2008 AGATATGCAGATGATACCACTTATGGCAGAAAGTGAAGAGAACTAAAGCCCTCTTG 2067
DB 121 AGATATGCAGATGATACCACTTATGGCAAAAGTGTAGAGAACTAAAGAGCCGCTTG 180
QY 2068 ATGAAGGTGAAGAGAGAGAGTGAAGAAAGTGGCTTAAAGCTCAACATTCAGAAAAAG 2127
DB 181 ATGAAGGTGAAGAGAGAGATGAAGAAAGTGGCTTAAAGTTCCTTAAATTCACATTC 240
QY 2128 ATCATGGCATCTGGT-CCCATCACTTCATGGGAATAGATGGGAAACAGTGGAAACAGT 2186
DB 241 ATCATAGCATCTGGTCCCATCATTTTCATGGCAAAATATATGGAGAAATTAATGAAACAG 300
QY 2187 GTCAGACTTTATTTTGGGGGGCTCCAAATCACTGCAGATGGTGAAGCCATGAA 2246
DB 301 GACAGATTTTATTTTCT--TGGGCTCCAAATCACTGTGATGGTGTGACCTGACCCCA 358
QY 2247 ATTAAGACACTTACTCTCTTGGAGAAAGTTA 2280
DB 359 ATTAAGACACTTATCTTGGAGAAAGCTA 392

RESULT 29
US-09-960-352-9592
; Sequence 9592, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.

Matches	337;	Conservative	0;	Mismatches	70;	Indels	1;	Gaps	1;
QY	2377	TGTGAGAGTTGGAC	TGTGCAAGAAAGCTGAC	CACTGAAGAATTGATGCTTTGAACTGTGG	2436				
DB	407	TGTGAGAGTTGGAGTAT	-AAGAAAGTTGACACCA	AGAAATGATGCTTTAAACTGTGG	349				
QY	2437	TGTTGGAGAGACT	TTTGAGAGTCCCTTGGAGCTGCAAGGAGATCCAA	CCAGTCCATTTCTG	2496				
DB	348	TGTTGAAAGAGACT	TTTGATAGTCCATGGGACTGCAAGAAAGATCAG	ACCAGTCAAAACCTA	289				
QY	2497	AAGGAGATCAGC	CCCTGGGATTTCTTTGGRAGGAATGATGCTAAAGCTG	AAACTCCAGTAC	2556				
DB	288	AATGAATCAGT	CCTTAATATTCATTGGAAGGACTGATGCTGAAGCTG	AAACTCCAGTAC	229				
QY	2557	TTTGGCCACCTGAT	CAGAAAGCTGACTCACTGGAAGAACCCCTGATGCTGGGAGGGATT	2616					
DB	228	TTTGGCCACCTGAT	TGAAGAACTGACTCATTTGAAAGAACCTGATGCTGGGAAAGATT	169					
QY	2617	GGGGCAGGAGGAG	AGGGGACACAGAGATGAGATGGCTGATGGCATC	ACTGACTCG	2676				
DB	168	GAAAGCATAAGG	AGAACCGGACACAGAGATGTGTTGAATGGCAT	CACCGACTCG	109				
QY	2677	ATGGACGTGAGT	CTGGGTGAACCTCTGGAGTTGGTGATGGACAGGAGGGCTGCTCTGCG	2736					
DB	108	ATGGACATGAGT	TGAGCAACTCCGGAGTTGGTGATGTACAGGGCAGCCTGGTATGCT	49					
QY	2737	GCAGTTCATGGGT	CAAGAGTTGGACACGACTGAGCAACTGAACT	2784					
DB	48	GGAGTCCATGGGT	CAAGAGTCTTAGACGACTGAGTCACTGAACT	1					

RESULT 31

US-09-960-352-13323

Sequence 13323, Application US/09960352

Patent No. US20020137139A1

GENERAL INFORMATION:

APPLICANT: Warren, Wesley C.

APPLICANT: Tao, Ningbing

APPLICANT: Byatt, John C.

APPLICANT: Mathialagan, Nagappan

TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND

TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION

FILE REFERENCE: 16511.006/37-21(10298)/C

CURRENT APPLICATION NUMBER: US/09/960,352

CURRENT FILING DATE: 2001-09-24

NUMBER OF SEQ ID NOS: 15112

SEQ ID NO 13323

LENGTH: 423

TYPE: DNA

ORGANISM: Bos taurus

OTHER INFORMATION: Clone ID: 57-LIB3057-008-Q1-K1-G2

US-09-960-352-13323

Query Match 7.9%; Score 282.9; DB 1; Length 423;

Best Local Similarity 83.0%; Pred. No. 1e+02;

Matches 356; Conservative 0; Mismatches 66; Indels 7; Gaps 3;

QY	1637	GACCAATTTGCCAA	CATCCTCTGTATCATGGAAGAAAGCAAGAGAGTTCCAGAAAAACATC	1696					
DB	1	GATCAATTTGCCAA	CATCTGTTGGATTATAGAAAAAGCAAGGGAATTCAGAAAAACATG	60					
QY	1697	TATTTCTGCTTTAT	TGACTATGCAAAAGCCTTTTGACTGTGGGGGTGCAATAAACTGTGG	1756					
DB	61	TACATCTACTTCA	TGAACCTACCTTTATAGCCTTTGACTGTGTGGATCACAACAACAC	-TGG 118					
QY	1757	AAAAATTTCAAAGG	ATGGGAATACCAAGACCACTTGACTGACTCTTGAAAAATTTGTAT	1816					
DB	119	AAAAATTTCAAAGG	ATGGGAATACCAAGACCACTTGACTGCTCTGAGAAATCTGTAT	178					
QY	1817	GCAGGTGAGGAAG	CAACAGTTAGAACTGGACATGGAAACAACAGACTGGTTCCAAGTAGGA	1876					
DB	179	GCAGGTCAAGAAG	AAAC-----AGAACTGGACATGGAAAAATGCACTGGTTCAAAATTGGG	234					

OTHER INFORMATION: Clone ID: 29-LIB3058-049-Q1-K1-H1
US-09-960-352-6765

Query Match 7.7%; Score 273.7; DB 1; Length 430;
Best Local Similarity 81.2%; Pred. No. 1.1e+02;
Matches 349; Conservative 0; Mismatches 48; Indels 33

[illegible]

RESULT 35

US-09-960-352-9318 ; Sequence 9318, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:

GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.

APPLICANT: Mathialagan, Nagappan

INVENTOR: Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOL
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION

FILE REFERENCE: 16511.006/37-21(10298)C

CURRENT APPLICATION NUMBER: US/0

;; CURRENT FILING DATE: 2001-

NUMBER OF SEQ

NUMBER OF SEQ
; SEQ ID NO 9318

LENGTH: 43

LENGTH: 430
TYPE: DNA

ORGANISM: Bos taurus

ORGANISM: BOB
OTHER INFORMATION:

Query Match 7.6%; Score 271.8; DB 1; Length 436;
Best Local Similarity 81.8%; Pred. No. 1.1e+02;
Matches 336; Conservative 0; Mismatches 67; Indels 8; Gaps 2;

1755	GGAAATTCCTGAAGGGATGGGAATACAGACAACCTGACTGACTCTGTGAAAAATTGT	1814	
	33	GGAAATTCCTTAAGAGATGGAAACACTAGACCACTTACTCTGGTCTCTGAGAAACCTGT	92
1815	ATGCAGTTCAGGAAGCAACAGTTAGAACTGGACATGGAAACAACAGACTGTGTCCAACTAG	1874	
	93	ATTTCAGTCAAGATGCAACAGTTAGAACTAGACATAGACCACCAACGCACCTGTGTCCAAATTG	152

Qy	1875	GA	AAAGAGGATGTC	CAAGGCTGTATAT	TGTGACCCGGCTTTGTTAACTTCTATGCAAG- 19333
Db	153	AG	AAAGAGGACTT	CAAGGCAGTGTAT	TGTGACCCCTGCTTATGTAACCTTATATGCAAGC 212
Qy	1934	AC	ATCATGAAACGCT	TGGGCTGGAAGAAGCA	CAAGCTGGAATCAAGATTGCCGGAGAA 1993
Db	213	AC	ATCATGTGAATG	CCAGTTGGATTAAGACA	CAAGTTGGAGTCAAGATTTCAAGGAGAA 272
Qy	1994	AT	AGCAATAACCTC	AGATATGCAAGATGATA	CCACCCTTATGGCAGAAAGTGAAGAGGAAC 2053
Db	273	AT	ATCAGTAACCTC	AGATATGCAAGAGAC	CACCATCCTTATG-----ATGAAGAGGAAC 325
Qy	2054	TA	AAAAAGCCTCTT	GATGAAGGTGAAGAGG	AGAGTGAATAAAGTTGGCTTTAAAGCTCAACA 2113
Db	326	TA	AAAGGCCTCTT	GATGAAGTGAATAAAG	GAGAGCAAAAAGCTGGCTTTAAAACTCAACA 385
Qy	2114	TT	CAGAAAAACGAAG	ATCATGGCATCTGGT	CCCATCACTTCATGGGAAATAG 2164
Db	386	TT	CAAAAAGTTAAG	ATCATGGCATCAAGGC	CATCACTTCATGGCAAGAG 436

RESULT 36

US-09-960-352-11236/c
; Sequence 11236, Application US/09960352
; Patent No. US20020137139A1

; GENERAL INFORMATION:

APPLICANT: Warren, Wesley C.

APPLICANT: Tao, Nengbing

APPLICANT: Byatt, John C.

APPLICANT: Machialagan, Nagapan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION

FILE REFERENCE: 16511.006/37-21 (10298)C

US/09/960,352

; CURRENT FILING DATE: 2001-

; COMMENT: LINE
; NUMBER OF SEQ; NUMBER OF SE
; SEQ ID NO 112

; LENGTH: 30

LENGTH: 320
TYPE: DNA

1 LIFE: DNA
2 ORGANISM: Bos taurus

ORGANISM: BOS C
; OTHER INFORMATION:

Query Match 7.6%; Score 270.2; DB 1; Length 396;
Best Local Similarity 84.5%; Pred. No. 1.2e+02;
Matches 326; Conservative 0; Mismatches 58; Indels 2; Gaps 2;

1018	QY	TTTGGCCTTGGAAATACGGAAATGAAGCAGGCGAAAGACTAAATAGAGTTTTTGCACAGAAAAT	1077
395	Db	TTTGGCCTTGGAGTACAAAATGAAGCAGGCGAAAGGCTAAACAGAGTTTTTGCACAAAGAAC	336
1078	QY	GCATCGTGTATAGCAAAACACCCCTCTTCCAAACAACAAGAGAAGACTCTACACATGGAACA	1137
335	Db	ACATTGGTGTATAGCAAAACATCCCTCTTGTCAACAACAACAATAGACAACATTTTACACATGAACA	276
1138	QY	TCACCAGATGGTCAACACCCGAATCAGATTGATTATATTTCTTTCCAGCCAAAGATGGAGA	1197
275	Db	TCATCAGATGTTCAATACCGATACAGACTGATTATATTTCTTTACAGTTGAAGATGAAGA	216
1198	QY	AGCTCTATACAGTCAGACAAAAACAAGCCAGAGAGCTTACTGTGGCTCAGATCATGAATC	1257
215	Db	AGCTCTATA-AGTTAGCAAAAACAAGACTGGGAGTTTACTGTGGCTCAGATCATGAATC	157
1258	QY	CTTATTGCCAAATTCAGACTTAAATTGAAGAAGTAGGGGAAAAACACTAGATCACTCAGG	1317
156	Db	CTTATTGCCAAATTCGAGACTT-AACTGAAGAAGATAGAGAAAAACCAAGAACTCATTCAG	98
1318	QY	TNAGACCTAAATCCATCCCTTATGATTATACAGTGGNAGTGCAGAAATAGATTTTAGGGC	1377
97	Db	TGTGACTTAAATCAAAATCCCTTACAATTTATACAGTGAAGTGCACAAACAGATTTCAGGGA	38
1378	QY	CTAGATCTGATAGACAGAGTACCTAA	1403

Db 37 TTAGATCTGATCAACAGAGTGTGAA 12
RESULT 37
US-09-960-352-7912
; Sequence 7912, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 7912
; LENGTH: 449
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 34-LIB3058-004-Q1-K1-A6
US-09-960-352-7912
Query Match 7.6%; Score 270.2; DB 1; Length 449;
Best Local Similarity 82.1%; Pred. No. 1.1e+02;
Matches 366; Conservative 0; Mismatches 58; Indels 22; Gaps 5;
QY 1666 GGAAAAAGCAGAGAGTCCAGAAAAACATCTATTTCTGTTTATTTGACTATGCAAAAGC 1725
Db 12 GGAAAAAGCAGAGAGTCCAGAAAAACATCTATTTCTGTTTATTTGACTATGCAAAAGC 71
QY 1726 CTTTGTACTGTGGGGTCCACATAAAGTGTGAAAAATTTCTGAAAGGGATGGGAATACAGA 1785
Db 72 TTTTATAT---GGATCACAACTGTGAAAAATCTTCTGTTTATTTGACTATGCAAAAGC 127
QY 1786 CCACCTGACCTGACTCTTGAATAATTTGATGCAAGTCCAGAGCAACAGTTAGAACTGG 1845
Db 128 CCACCTTACCTGACAC---AGAAATCTGTGTGAGGCAAGAAACACAGTTAGTACTGG 184
QY 1846 ACATGGAAACACAGAGCTGGTTCCAGTAGGAAAAAGGAG-----TATGTCAG 1892
Db 185 ACATGGAAACACAGAGTGTTCAAAATTTGGGAAGGAGTCAAGGACAAATATATGTCAAG 244
QY 1893 GCTGTATTTGTCACCGCTGTTTAACTTCTATGAGAG-ACATCATGAGAAACGCTG 1951
Db 245 GCTATATATTGTCACCTGCTTATTTAACTTATATGTCACAGTACATCATGAGAAATGCCA 304
QY 1952 GGCTGGAAGAGCAACAGCTGGAATCAAGATTGCGGGAGAAATAGCAATAACCTCAGAT 2011
Db 305 GACTGGATGAAGCAACAGCTGGAATCAAGATTGCTGGGAGAAATATCAATAACCTCAGAT 364
QY 2012 ATGCAGATGATACCAACCTTTATGCGAGAAAGTGAAGAGGAACTAAAAAGCTCTTTGATGA 2071
Db 365 ATGCAGATGATGCCA-CCTTATGTTAGAAAGTGAAGAGGAACTAAAGAGCTCTTTGATGA 423
QY 2072 AGGTGAAAGAGGAGTCAAAAAGTT 2097
Db 424 AGGTGAAAGAGGAGAGTCAAAAAGCT 449
RESULT 38
US-09-960-352-11744/c
; Sequence 11744, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND

; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 11744
; LENGTH: 347
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 50-LIB34-023-Q1-E1-E6
US-09-960-352-11744
Query Match 7.5%; Score 269.3; DB 1; Length 347;
Best Local Similarity 90.1%; Pred. No. 1.3e+02;
Matches 299; Conservative 0; Mismatches 32; Indels 1; Gaps 1;
QY 2470 GCAAGGAGATCCAAACAGTCCATTCTGAAGAGATCAGCCCTGGGATTTCTTTGGAAGA 2529
Db 346 GCAAGGCGATCCAAACAGTTCATTCTGAAGGAGATCAGCCCTGGGATTTCTTTGGAAGA 287
QY 2530 ATGATGCTAAAGCTGAAACTCCAGTACTTTGCCCACTGATCAGAGAGCTGACTCACTG 2589
Db 286 ATGATGCTAAAGCTGCAACTCCAG-ACTTGACCACTCATGCGAAGAGTTGACTCATTTG 228
QY 2590 GAAAAAGCCCTGATGCTGGGAGGGATTGGGGCAGGAGGAGAGGGGACGACAGAGGATG 2649
Db 227 GAAAAGACTCTGATGCTGGGAGGGACTGGGGCAGGAGGAGAGGGGACACAGGGGATG 168
QY 2650 AGATGGCTGGATGGATCACTGACTCTGATGAGTGTGAGTGTGGGTGAACCTCTGAGTTG 2709
Db 167 AGATGGCTGGATGGATCACTGACTCTGATGAGTGTGAGTGTGGGTGAACCTCTGAGTTG 108
QY 2710 GTGATGGACAGGAGGCGCTGCTCGGGCGATTCTGGGGTCAAAAGAGTTGGACACGA 2769
Db 107 GTGATGGACAGGAGGCGCTGCTGCTGTGATTCATGGGTGCGCAAGAGTGGACACGA 48
QY 2770 CTGAGCAACTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2801
Db 47 CTGAGCGCACTAAACTTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 16
RESULT 39
US-09-960-352-7893/c
; Sequence 7893, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 7893
; LENGTH: 383
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (311)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: 34-LIB3057-006-Q1-K1-A6
US-09-960-352-7893
Query Match 7.5%; Score 268.1; DB 1; Length 383;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 330; Conservative 0; Mismatches 50; Indels 5; Gaps 3;
QY 2001 TAACCTCAGATATGCAAGTATCAACCTTATGGCAGAAAGTGAAGAGAACTAAAAAG 2060

Db		383	TAACCTTCAGATATCCGAGATGACCACTTTATGTCAGAGAGTGAACAGCAACTACAGAG	324
Qy		2061	CCTCTTGATGAAAGGTGAAGAAGGAGAGTGA AAAAGTTGGCTTAAAGTC- AACATT CAGA	2119
Db		323	CCTCGGTATGATNGTGTATACAGTAGAGTGAGCAAAGTTGGCTTAAAGCTTGAACAATTCAGA	264
Qy		2120	AACCAAGATCATGGCATTCTGGTCCCACACTTCATGGGAAATAGATGGGGAACAGTGG	2179
Db		263	GCACACAGATCATGGCAATTCAGTCCGATCACTTCATGGGAAATAGATGGGGAACAAATGG	204
Qy		2180	AAACAGTGT CAGACTTTATTTTTTGGGGGGCTCCAAAATCACTGCAGATGGTGA CTGCGAG	2239
Db		203	AAACAGTGT CATACTTTATGTTTT - TGGGGTCCAAAAATCACTGCAGATGGTGA CTGCGAG	146
Qy		2240	CCATGAAATTTAAAGACACACTTACTCCTTTGGAAGAAAAAGTTA-- ACCAACCTTAGATAGCAT	2297
Db		145	CCATGAAAGTGACAGACACTTACTCCTTTGGAAGGANAAGTTATGTCCTCAACG TAGATAGCAT	86
Qy		2298	ATTGARAAGCAGAGACATTACCCTTGCCAA CA AAGCCCCATCTAGTCAAGCCTATGGTTTT	2357
Db		85	ATTTAAANTCAGAGATAATGCTTTACCATCA AAGGTCGGTTAGTCAAGGCTATGGTTTT	26
Qy		2358	TCCAGTGGTCA TGATGATGTGAG	2382
Db		25	TCCAGTGGTCA TATATGATGTGAG	1

RESULT 40

```

US-09-960-352-2913/C
; Sequence 2913, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Ningbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND FAT DEPOSITION
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO. 2913
; LENGTH: 432
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 13-LIB3058-024-Q1-K1-D1
US-09-960-352-2913

```

Qy	2345	AGG	TATG	CTTTT	CCAGT	GGT	GTATG	2374	
Db	134	AGG	TG	GGT	TTTTT	CCAGT	GGT	CATATAAG	105
RESULT 41									
US-09-876-143-32/c									
; Sequence 32, Application US/09876143									
; Publication No. US20040081958A1									
; GENERAL INFORMATION:									
; APPLICANT: Infigen Inc.									
; APPLICANT: EILERTSEN, KENNETH J.									
; APPLICANT: PFISTER-GENSKOW, MARTHA									
; APPLICANT: CHILDS, LYNETTE									
; APPLICANT: FORSYTHE, TODD									
; APPLICANT: BISHOP, MICHAEL D.									
; TITLE OF INVENTION: IDENTIFICATION AND USE OF MOLECULAR MARKERS INDICATING									
; TITLE OF INVENTION: CELLULAR REPROGRAMMING									
; FILE REFERENCE: 028040-0202									
; CURRENT APPLICATION NUMBER: US/09/876,143									
; CURRENT FILING DATE: 2001-06-06									
; PRIOR APPLICATION NUMBER: 60/209,874									
; PRIOR FILING DATE: 2000-06-07									
; NUMBER OF SEQ ID NOS: 1744									
; SOFTWARE: PatentIn version 3.0									
; SEQ ID NO 32									
; LENGTH: 347									
; TYPE: DNA									
; ORGANISM: Bovine									
; FEATURE:									
; NAME/KEY: misc feature									
; LOCATION: (1)..(347)									
; OTHER INFORMATION: n is a, c, g, or t									
US-09-876-143-32									
Query Match									
Best Local Similarity 7.4%; Score 266; DB 1; Length 347;									
Matches 302; Conservative 0; Mismatches 27; Indels 2; Gaps 2;									
Qy	1799	CTCTTGAA	AAATTTG	TATG	CAGT	CGAGG	AAAGAG	TATG	CTCAAGGCTGTATATTGTACACCGCGCTTGTTT
Db	331	CTCTTGAG	AATCGT	GTATG	CAGT	CGAGG	CGCAAC	AGTAA	GACCTGGACATGAAGCAACA
Qy	1859	GACTG	GTTCC	AACTG	AGAAAG	GAGATG	ATG	CTCAAGGCTGTATATTGTACACCGCGCTTGTTT	1918
Db	271	GCGT	GTTCC	AAATG	AGAAAG	GAGTAC	GTG	CTCAAGGCTGTATATTGTACACCGCTTGTTT	212
Qy	1919	AACTT	CTATG	CAGAG	-ACAT	CATG	AGAAAC	CGCTGGGCTGGGAGCAACAAGCTGGAATC	1977
Db	211	ATCTT	ATATG	CAGAGT	ACAT	CATG	AGAAAC	CGCTGGGCTGGGAGCAACAAGCTGGAATC	152
Qy	1978	AGATT	GCCGG	GAGAA	TAGCA	ATAAC	CTC	CAGTATG	TCAGATGAT-ACACCCCTTATGGC
Db	151	AGATT	GCCAG	GAGAA	TAAT	ATCA	ATAAC	CTC	CATATATG
Qy	2037	AGAA	GCTGA	AGAG	GAAC	TAA	AAAG	CCCTCTT	GTATGAAGCTGAAAGAGGAGTCAAAAAAGT
Db	91	AGAA	AGCTG	AGAG	GAAC	TAA	AAAG	CCCTCTT	GTATGAAGCTGAAAGAGGAGGAAAAAGT
Qy	2097	TG	CTTTAA	AGCTCA	ACATT	CA	AGAA	ACGAAG	2127
Db	31	TG	CTTTAA	AGCTCA	ACATT	CA	AGAA	ACGAAG	1

; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 14979
; LENGTH: 423
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 64-LIB3058-007-Q1-K1-H8
US-09-960-352-14979

Query Match 7.2%; Score 258.1; DB 1; Length 423;
Best Local Similarity 91.9%; Pred. No. 1.3e+02;
Matches 283; Conservative 0; Mismatches 24; Indels 1; Gaps 1;

QY 1589 GAATCTCCAGTGTTCAGAGCTGTTTAGAAAAAGTCAGAGAACCCAGAGACCAAAATGGC 1648
DB 320 GAATCTCCAGTGTTCAGAGCTGTTTAGAAAAAGTCAGAGAACCCAGAGATCAAAATGGC 261
QY 1649 AAGATCTCTGTATCATGGAAGAACAGAGAGTTCAGAGAAACATCTATTTCTGCTTT 1708
DB 260 AAGATCTCTGTATCATGGAAGAACAGAGAGTTCAGAGAAACATCTATTTCTGCTTT 201
QY 1709 ATTGACTATGCAAGAGCTTTGACTGTGGGGTCAAAAT-AAAGTGTGGAATTTCTGAA 1767
DB 200 ACTGACTTTGCCAAGTCTTTGACTGTGTGATGATCAATTAAGTGTGGAATTTCTGAA 141
QY 1768 AGGATGGGAATACCAAGACCTGACTGACTGTTGAAAAATTTGTATCAGGTGAGGA 1827
DB 140 AGAGATGGGAATACCAAGACCTGACTGACTGTTGAAAAATTTGTATCAGGTGAGGA 81
QY 1828 AGCAACAGTGTAGACTGCAATGGAACACAGAGCTGTTTCAAGTAGGAAAAAGAGTATG 1887
DB 80 AGCAACAGTGTAGACTGCAATGGAACACAGAGCTGTTTCAAGTAGGAAAAAGAGTATG 21
QY 1888 TCAAGGCT 1895
DB 20 TCAAGGCT 13

RESULT 43
US-09-960-352-149/c
; Sequence 149, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 149
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 01-LIB34-016-Q1-E1-A1
US-09-960-352-149

Query Match 7.2%; Score 257.8; DB 1; Length 417;
Best Local Similarity 82.1%; Pred. No. 1.3e+02;
Matches 330; Conservative 0; Mismatches 62; Indels 10; Gaps 3;

QY 1854 CACAGACTGTTTCAAGTAGGAAAGAGTATGTCAGAGCTGTATATGTCACCG-GC 1912
DB 393 CAATGACTGTTTCAAAATTTGGAAAGGAGTACCAAGAGGTGTATATGTCACCTTAGC 334
QY 1913 TTGTTTAACTTCTATGCAGAGATCATGAGAAAAGCTGGGTGGAAGAGCAACAAGCTG 1972

DB 333 TTATTTAACTTGAAT-----ACATCAGGCAAAATGCCGAGTGGATGAACCAAGCTG 280
QY 1973 GAATCAAGATTGCCGGGAGAAATAGCATAAATCTCAGATATGCAGATGATACCAACCTTA 2032
DB 279 GAGTCAAGATTGCCGGGAGACATATCAACAATCTCAGATATGTAGATTAATACCACTTAA 220
QY 2033 TGGCAGAAAAGTGAAGAGAACTAAAAAGCCTCTTGTGAAGAGTGAAGAGGAGAGTGA 2092
DB 219 TGGCAGAAAAGCAAGAGAACTAAAGAGCTATTGTGAAGTGAAGAGGAGAAAGTGA 160
QY 2093 AAGTTGGCTTAAAGCTCAACATTCAGAAAAAGAGATCATGCACTGTGTCCTCATCACTT 2152
DB 159 AAGCTGGCATAAACTCAAAATTTCAAAAAGTGAATCTGCACTGTGCACTCATCACTT 100
QY 2153 CATGGAAATAGATGGGGAACAGTGGAAACAGTGTGCACTTATTTTGGGGGGCTC 2212
DB 99 CATGCCAATAGATGGGGAACAGTGGAAATGGAAACAGTGGCAATTTTATTTCTT---GGGCTC 43
QY 2213 CAAAATCACTGCAGATGTGACTGCAGCCATGAAATTTAAAG 2254
DB 42 CAAAATCACTGCAGACAGTGAATGAGCCATGAAATTTAAAG 1

RESULT 44
US-09-960-352-7209
; Sequence 7209, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 7209
; LENGTH: 456
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 31-LIB3058-021-Q1-K1-H11
US-09-960-352-7209

Query Match 7.1%; Score 254.8; DB 1; Length 456;
Best Local Similarity 83.6%; Pred. No. 1.3e+02;
Matches 300; Conservative 0; Mismatches 57; Indels 2; Gaps 1;

QY 2431 CTGTGTTTGGAGAGACTCTTGAGAGTCCCTTGACTGCAAGGAGATCCCAACAGTCC 2490
DB 1 CTGTGTTTGGAGAGACTCTTGAGAGTCCCTTGACTGCAAGGAGATCCCAACAGTCC 60
QY 2491 ATTCTGAAGGAGATCAGCCCTGGGATTTCTTTGGAGGAATGATCTAAAGTGAACCTC 2550
DB 61 ATCTTAAAGAAATCAATCTGAATATTCACTGGAAGGAGTGTGCTGAAGCTGAACTC 120
QY 2551 CAGTACTTTGGCCACCTGTATCAGAGAGTCACTCACTGGAAGAGCCCTGTGCTGGGA 2610
DB 121 CAGTACTTTGGCCACCTGTATGCAAGAACTGACTCTCTTTGAAAGAGCCCTGTGCTGGG 180
QY 2611 GGGATTGGGGGAGGAGGAGAGGGGACACAGAGGATGAGATGGCTGGATGCATCACT 2670
DB 181 AAGATTGAAGGAGGAGGAGAGGGGACACAAAGGATGAGATGGTTGGATGCATCACC 240
QY 2671 GACTCGATGAGCTGTAGTCTGGGTGAATCTCTGGAGTTGGTGTATGACAGGGAGGCTGT 2730
DB 241 GATTCAATGGACATGAGTTGGGTAACTCTGGAGCTGGAGATGACAGGGAGGCTGG 300
QY 2731 CTTGGCGGAGTTTCAATGGGTTCACAAAGAGTGGACACAGTGAACCTGAACCTGAAC 2789
DB 301 CGTGCTACAGTCCATGAGGTGCG--AGGCCGAGACATGACTGAGAGCTGAACCTGAAT 357

Search completed: August 24, 2004, 13:39:08
Job time : 285 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 13:40:03 ; Search time 73 Seconds
(without alignments)
3.947 Million cell updates/sec

Title: US-10-664-775-2
Perfect score: 3572
Sequence: 1 gtcaggagggcgagctga.....gcaacacagcagaagctt 3572

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 0.5

Searched: 45 seqs, 40327 residues

Total number of hits satisfying chosen parameters: 90

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : rst2.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	829.8	23.2	909	1	CC583446
C 2	824.6	23.1	926	1	ACCESSION:CC583446
C 3	824.3	23.1	949	1	ACCESSION:CC479701
C 4	821.9	23.0	927	1	ACCESSION:CC479360
C 5	818.2	22.9	894	1	ACCESSION:CC508172
C 6	810.2	22.7	919	1	ACCESSION:CC544978
C 7	806.8	22.6	882	1	ACCESSION:CC577078
C 8	801.7	22.4	866	1	ACCESSION:CC514626
C 9	800.3	22.4	887	1	ACCESSION:CC586882
C 10	800.1	22.4	951	1	ACCESSION:CC520949
C 11	798.2	22.3	936	1	ACCESSION:CC548437
C 12	798.2	22.3	938	1	ACCESSION:CC594700
C 13	797.8	22.3	922	1	ACCESSION:B2882580
C 14	797.2	22.3	899	1	ACCESSION:CC480115
C 15	796.9	22.3	958	1	ACCESSION:CC475230
C 16	796.001	22.3	944	1	ACCESSION:CC593631
C 17	793.1	22.2	914	1	ACCESSION:CC503865
C 18	790.8	22.1	904	1	ACCESSION:CC536522
C 19	790.2	22.1	897	1	ACCESSION:CC476492
C 20	785.2	22.0	881	1	ACCESSION:CC579657
C 21	784.6	22.0	925	1	ACCESSION:CC500971
C 22	784.2	22.0	934	1	ACCESSION:CC543667
C 23	783.8	21.9	907	1	ACCESSION:CC547247
C 24	782.8	21.9	907	1	ACCESSION:CC480290
C 25	779.601	21.8	949	1	ACCESSION:CC589137
C 26	779.1	21.8	890	1	ACCESSION:CC541222
C 27	778.3	21.8	881	1	ACCESSION:CC546101
C 28	778.1	21.8	945	1	ACCESSION:CC511457
C 29	775.4	21.7	833	1	ACCESSION:CC495104
C 30	774.7	21.7	887	1	ACCESSION:CC539858
C 31	774.3	21.7	896	1	ACCESSION:CC528566
C 32	774.2	21.7	827	1	ACCESSION:CC508740
C 33	774.2	21.7	898	1	ACCESSION:CC512039

Query Match

23.2%; Score 829.8; DB 1; Length 909;

RESULT 1
CC583446/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

ALIGNMENTS

34	773.7	21.7	881	1	CC581072
C 35	771.3	21.6	915	1	ACCESSION:CC536979
C 36	770.7	21.6	846	1	ACCESSION:CC505149
C 37	769.501	21.5	947	1	ACCESSION:CC591633
C 38	768.8	21.5	858	1	ACCESSION:CC586052
C 39	768.5	21.5	877	1	ACCESSION:CC525458
C 40	768.4	21.5	841	1	ACCESSION:CC548709
C 41	768.4	21.5	865	1	ACCESSION:CC533944
C 42	767.3	21.5	822	1	ACCESSION:CC581625
C 43	767	21.5	861	1	ACCESSION:CC595021
C 44	766.4	21.5	870	1	ACCESSION:CC512505
C 45	765.9	21.4	862	1	ACCESSION:CC540150

CC583446 909 bp DNA linear GSS 18-JUN-2003
CH240_379F21.T7 CHORI-240 Bos taurus genomic clone CH240_379F21,
genomic survey sequence.

CC583446 GI:31933114

GSS.

Bos taurus (cow)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovidae; Bovinae; Bos.

1 (bases 1 to 909)

Holt,R., Scott,J., Yang,G., Barber,S., Smalilus,D., Prabhu,A.-L.,

Tsai,M., Cloutier,A., Lee,D., Ginn,N., Olson,T., Mayo,M., Chiu,R.,

Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R.,

Mathewson,C., Wye,N., Masson,A., Brown-John,M., Jones,S.,

Schein,J., Marra,M., de Jong,P., McWilliam,S., Barris,W.,

Dalrymple,B.P. and Tellam,R.

Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398

Unpublished (2003)

Other_GSSs: CH240_379F21.TABAC13P2

Contact: Rob Holt

Sequencing

The British Columbia Cancer Agency Genome Science Centre

600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6

Tel: 604-877-6085

Fax: 604-877-8276

Email: rholt@bccgsc.ca

Clones are derived from the bovine BAC library CHORI-240

(http://www.chori.org/bacpac/bovine240.htm). For BAC library

availability, please contact Pieter de Jong (pdejong@mail.cho.org).

Clones may be purchased from BACPAC Resources

(http://www.chori.org/bacpac/ordering-information.htm). This work

was undertaken as part of the International Bovine BAC Mapping

Consortium (IBMC) by CSIRO Livestock Industries, Australia and the

British Columbia Genome Sciences Centre, Canada.

Plate: 379 row: F column: 21

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

1..909

/organism="Bos taurus"

/mol_type="genomic DNA"

/strain="Breed: Hereford"

/db_xref="taxon:9913"

/clone="CH240_379F21"

/sex="Male"

/cell_type="Blood"

/clone_lib="CHORI-240"

/notes="Vector: pTABAC1.3; Site: 1: MboI; Site 2: MboI;

Hereford bull li Domino 99375; CHORI-240 Bovine BAC

library (Male) produced by Pieter de Jong"

Best Local Similarity 95.3%; Pred. No. 0.11;
Matches 855; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 533 GATGCTGTAAGAGCAATATTCATAGGAAGCTGGAATGTGAGGTCATGATCAAGCA 592
Db 897 GATGATGTAAGAGCAATATTCATAGGAAGCTGGAATGTGAGGTCATGATCAAGCA 838
QY 593 AATTGGAAGTGTCAAAACAAAGAGATGGAAGAGTCAAGTCAACATCTAGGAATCAGCG 652
Db 837 AATTGGAAGTGTCAAAACAAAGAGATGGAAGAGTCAAGTCAACATCTAGGAATCAGCG 778
QY 653 AACTAAATGGACTCGAATGGGTGAATTTAACTCAGATGACCATTTATATCTACTACTGCG 712
Db 777 AACTGAAATGGACTCGAATGGGTGAATTTAACTCAGATGACCATTTATATCTACTACTTCA 718
QY 713 GGCAGGATCCCTCAGAGAAATGAGTAGGATGAGCATGCTGTCACAAAGAGTCCGAATG 772
Db 717 GGCAGGATCCCTCAGAGAAATGAGTAGGATGAGCATGCTGTCACAAAGAGTCCGAATG 658
QY 773 CAGTACTTGGATGCACTCTCAAAACGACAGAGATGATCTCTGTTGTTTCCAAAGGCAAC 832
Db 657 CAGTACTTGGATGCACTCTCAAAATGACAGATGATCTCTGTTGTTTCCAAAGGCAAC 598
QY 833 CATTCATATCACAGTAATCAAGTCTATGCCCCAACAGTAATGCTGGAAGACTGAAG 892
Db 597 CATTCATATCACAGTAATCAAGTCTATGCCCCAACAGTAATGCTGGAAGACTGAAG 538
QY 893 TTGAACGCTCTATGAAGACCTCAAGACCTTTTAGAACTAAACACCCCAAAAGAGATGTC 952
Db 537 TTGAACGCTCTATGAAGACCTCAAGACCTTTTAGAACTAAACACCCCAAAAGAGATGTC 478
QY 953 TTCTCATATAGGGAATGGAATGCAAAAGTAGGAAGCAAAACACCTGGAGTAACAG 1012
Db 477 TTCTCATATAGGGAATGGAATGCAAAAGTAGGAAGTAAGAAACACCTGGAGTAACAG 418
QY 1013 GCAAAATTTGGCTTTGGATACGGAATGAGGAGGCAAGACTAATAGATTTTGCCAG 1072
Db 417 GCAAAATTTGGCTTTGGATACGGAATGAGGAGGCAAGACTAATAGATTTTGCCAG 358
QY 1073 AAAATGCACTGGTATAGCAAAACCCCTTTTCCAAACCAACAGAGAAAGACTCTACACAT 1132
Db 357 AAAATGCACTGGTATAGCAAAACCCCTTTTCCAAACCAACAGAGAAAGACTCTACACAT 298
QY 1133 GGAATACACAGATGCTCAACACCGAATCAGATTGATTATTTCTTTGAGGCAAGAT 1192
Db 297 GGAATACACAGATGCTCAACACCGAATCAGATTGATTATTTCTTTGAGGCAAGAT 238
QY 1193 GGAGAACTCTATACAGTCAAGCAAAACCAAGACGAGCTTACTGTGCTCAGATCATG 1252
Db 237 GGAGAACTCTATACAGTCAAGCAAAACCAAGACGAGCTTACTGTGCTCAGATCATG 178
QY 1253 AACTCTTTATGCAAAATTCAGACTTAAATGGAAGAAAGTAGGAAACCACTAGATCAC 1312
Db 177 AACTCTTTATGCAAAATTCAGACTTAAATGGAAGAAAGTAGGAAACCACTAGATCAC 118
QY 1313 TCAGGTAAGACCTAAATCCATCCCTTATGATTATATACAGTGAAGTGAAGATAGATTTA 1372
Db 117 TCAGGTAAGACCTAAATCCATCCCTTATGATTATATACAGTGAAGTGAAGATAGATTTA 58
QY 1373 AGGCCCTAGATCTGATAGACAGATGCTTAATGAACCTATGACAGAGGTTTCATGACA 1429
Db 57 AGGCCCTAGATCTGATAGATAGTGCCTGATGAACCTATGGAATGAGGTTTCGTGACA 1

RESULT 2
LOCUS CC479701 926 bp DNA linear GSS 15-JUN-2003
DEFINITION CH240_306H6.T7 CHORI-240 Bos taurus genomic clone CH240_306H6,
genomic survey sequence.
ACCESSION CC479701
VERSION CC479701.1 GI:31757509
KEYWORDS GSS.
SOURCE Bos taurus (cow)

ORGANISM

Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactylia; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.

REFERENCE

1 (bases 1 to 926)

AUTHORS

Holt, R., Stott, J., Yang, G., Barber, S., Smalilus, D., Prabhu, A.-L.,
Tsai, M., Cloutier, A., Lee, D., Girn, N., Olson, T., Mayo, M.,
Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
Schein, C., Marra, M., de Jong, P., McWilliam, S., Barris, W.,
Dalrymple, B. P. and Tellam, R.

TITLE

Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398

JOURNAL

Unpublished (2003)

COMMENT

Other GSSs: CH240_306H6.TARBAC13P2

Contact: Rob Holt

Sequencing

The British Columbia Cancer Agency Genome Science Centre

600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6

Tel: 604-877-6085

Fax: 604-877-6276

Email: rholt@bccg.ca

Clones are derived from the bovine BAC library CHORI-240

(<http://www.chori.org/bacpac/bovine240.htm>). For BAC library

availability, please contact Pieter de Jong (pdejong@mail.cho.org).

Clones may be purchased from BACPAC Resources

(<http://www.chori.org/bacpac/orderinginformation.htm>). This work

was undertaken as part of the International Bovine BAC Mapping

Consortium (IBMC) by CSIRO Livestock Industries, Australia and the

British Columbia Genome Sciences Centre, Canada.

Plate: 306 row: H column: 6

Seq primer: T7

Class: BAC ends.

FEATURES

Location/Qualifiers

1..926

/organism="Bos taurus"

/mol_type="genomic DNA"

/strain="bred: Hereford"

/db_xref="taxon:9913"

/clone="CH240_306H6"

/sex="Male"

/cell_type="Blood"

/clone_lib="CHORI-240"

/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;

Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC

library (Male) produced by Pieter de Jong"

Query Match 23.1%; Score 824.6; DB 1; Length 926;

Best Local Similarity 94.4%; Pred. No. 0.11;

Matches 875; Conservative 0; Mismatches 50; Indels 2; Gaps 2;

QY 45 GGTAAAGGAGCAGTAGCTCGCTTTGCTGGAGCCGTTAAAGAGATACCCACGCCCAAG 104

Db 1 GGTAAAGGAGCAGTAGCTCGCTTTGCTGGAGCCGTTAAAGAGATACCTCACGCCCAAG 60

QY 105 GTAACAGAAACCCAGTAGTGGTGGTGTGTGAGAGGCATCAGAGGCACATAC 164

Db 61 GTAAGAGATACCCAGTAGTGGTGGTGTGTGAGAGGCATCAGAGGCACATAC 120

QY 165 TGAAACCATACACGAGAAACTAGTCAATCTAATCACTAGGACCAACGCTTTGCTA 224

Db 121 TGAAACCATACCTCAGAGAAACTAGTCAATCTAATCACTAGGACCAACGCTTTGCTA 180

QY 225 ACTCAATCAAACTAAGCCATGCCGTGGGCAACCCAGATGGGAGGTATGTTGAGA 284

Db 181 ACTCAATCAAACTAAGCCATGCCGTGGGCAACCCAGATGGGAGGTATGTTGAGA 240

QY 285 GATCTGACAGAACTGTGGTCCACTGCGAGAGGGAAT-GCAAAACCACTTCAGTATTCCTGCC 343

Db 241 GATCTGACAGAACTGTGGTCCACTGCGAGAGGGAATGGAACCACTTCAGTATTCCTGCC 300

QY 344 TTGAGAACCCCATGAACAGTATGAAAAGGCAAAATGATAGTACTGAAGAGGAATCC 403

Db 301 TTGAGAACCCCATGAACAGTATGAAAAGGCAAAATGATAGTACTGAAGAGGAATCC 360

(http://www.chori.org/bacpac/bovine240.htm). For BAC library availability, please contact Pieter de Jong (pdejong@mail.choi.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/ordering_information.htm). This work was undertaken as part of the International Bovine BAC Mapping Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the British Columbia Genome Sciences Centre, Canada.

Plate: 306 row: K column: 22

Seq primer: 17

Class: BAC ends.

Location/Qualifiers

source

1..949

/organism="Bos taurus"

/mol_type="genomic DNA"

/strain="breed: Hereford"

/db_xref="taxon:9913"

/clone="CH240_306K22"

/sex="Male"

/cell_type="Blood"

/clone_lib="CHORI-240"

/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI; Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC library (Male) produced by Pieter de Jong"

Query Match 23.1%; Score 824.3; DB 1; Length 949;
Best Local Similarity 94.2%; Pred. No. 0.11;
Matches 896; Conservative 0; Mismatches 48; Indels 7; Gaps 4;

1667 GAAAAGCAAGAGAGTTCCAGAAACATCTATTCTGCTTTTATTGACTATCCAAAAGCC 1726
Db 3 GAAAAGCAAGAGAGTTAAAGAAACATCTATTCTGCTTTTATTGACTATCCAAAAGCC 62
1727 TTTGACTGTGGGGTCCAAATAAATCTGGAATAATCTGAAAGGGATGGAAATACAGAC 1786
Db 63 TTTGACTGTGGATCCAAATAAATCTGGAATAATCTGGAATAATCCAGAC 120
1787 CACCTGACCTGACTCTTGAATAATTTGATGAGTCCAGGACCAACAGTTAGAACTGCA 1846
Db 121 CACCTGACCTGCTCTTGAGAAACCTGTATGAGGTCCAGGACCAACAGTTAGAACTGCA 180
1847 CATGGAACCAACAGACTGTTCCAAAGTAGGAAAGAGATGTCAAGGCTGTATATTGTCA 1906
Db 181 CATGGAACCAACAGACTGTTCCAAAGTAGGAAAGAGATGTCAAGGCTGTATATTGTCA 240
1907 CCGGCTTTTAACTTTATGACAGAG-ACATCATGAGAAACGCTGGCTGGAAGAGCA 1965
Db 241 CCGGCTTTTAACTTTATGACAGAGTACATTAAGTCAAGGCTGGAAGAGCA 300
1966 CAGCTGGAATCAAGATTCGCGGAGATAGCAATAACCTCAGATATCAGATGATAC 2025
Db 301 CAGCTGGAATCAAGATTCGCGGAGATAGCAATAACCTCAGATATCAGATGATAC 360
2026 ACCCTTTATGGCAGAAAGTAGGAGGAACTTAAAGGCTCTTTCATGAAGTGAAGAGGAG 2085
Db 361 ACCCTTTATGGCAGAAAGTAGGAGGAACTTAAAGGCTCTTTCATGAAGTGAAGAGGAG 420
2086 AGTGAAGAAAGTTGGCTTTAAAGCTCAACATTCAGAAACCAAGATCATGCATCTGCTCC 2145
Db 421 AGTGAAGAAAGTTGGCTTTAAAGCTCAACATTCAGAAACCAAGATCATGCATCTGCTCC 480
2146 ATCACTTCATGGAAATAGATGGGAAACAGATGGAACAGTGTCAAGCTTTATTTTGG 2205
Db 481 ATCACTTCATGGAAATAGATGGGAAACAGTGTGAACAGTGTCAAGCTTTATTTT 538
2206 GGGGCTCCAAATCACTGACAGATGGTGAAGTCCAGCCATGAATTAAGACACTTACTCC 2265
Db 539 GGGGCTCCAAATCACTGACAGATGGTGAAGTCCAGCCATGAATTAAGACACTTACTCC 598
2266 TTGGAAGAAAGTTA--ACCACTTAGATAGCATATTGAAAGCAGAGACATTACCTTGC 2323
Db 599 TTGGAAGAAAGTTATGACCACTAGATAGCATATTCAAAGCAGAGACATTACTTGT 558
2324 CAACAAGGCCCATCTAGTCAAGGCTATGTTTTTCCAGTGGTCAATGTATGATGTGAGA 2383

404 CCAGGTCAAGTGGTGGCCCATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGA 463
Db 361 CCAGGTCAAGTGGTGGCCCATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGA 420
464 ATGAAGAGATGGAGCCCAACAAAGAAATACCCAGCTGGAGTGGATGACTGGTGATATAA 523
Db 421 AAGAGGATGGAGCCCAACAAAGAAATACCTCAGCTGGAGTGGATGACTGGTGATATAA 480
524 GCAAGGTCGATGCTGTAAGAGCAATATTGCTATAGAACTCGAATGTCCAGTCCATGA 583
Db 481 GCAAGGTCGATGCTGTAAGAGCAATATTGCTATAGAACTCGAATGTCCAGTCCATGA 540
584 ATCAAGCAATGGAGTGGTCAACAAAGAGATGGCAAGTGAATGCAATCTAG 643
Db 541 ATCAAGCAATGGAGTGGTCAACAAAGAGATGGCAAGTGAATGCAATCTAG 600
644 GAATCAGCAACTAAATGACTGGAATGGTGAATTTAACTCAGATGACCATATATCT 703
Db 601 GAATCAGCAACTGAATGACTGGAATGGTGAATTTAACTCAGATGACCATATATCT 660
704 ACTACTCGGGCAGGAATCCCTCAGAAGAAATGGAGTAGCCATCATGTCACAAAGAG 763
Db 661 ACTACTCGGGCAGGAATCCCTCAGAAGAAATGGAGTAGCCATCATGTCACAAAGAG 720
764 TCCGAATGAGTACTTGGATGCTCTCAAAACAGCAGATGATCTCTGTTTGTTC 823
Db 721 ACCGAATGAGTACTTGGATGCTCTCAAAACAGCAGATGATCTCTGTTTGTTC 780
824 AAGCAAAACCATCAATATCAGATTAATCCAAAGTCTATGCCCCAACAGTAATGCTGAAG 883
Db 781 AAGCAAAACCATCAATATCAGATTAATCANAGTCTATGCCCCAACAGTAATGCTGAAG 840
884 AAGTGAAGTGAAGTCTCTTGAAGCTTACAGACCTTATGAACTTATGAACTACACCCAAA 943
Db 841 AAGTGAAGTGAAGTCTCTTGAAGCTTATGAACTTATGAACTTATGAACTACACCC-AAA 899
944 AAGTGTCTCTCTCATTTATAGGGACT 970
Db 900 AAGTGTCTCTCTCATTTATAGGGAT 926

CC479360 949 bp DNA linear GSS 16-JUN-2003
CH240_306K22.T7 CHORI-240 Bos taurus genomic clone CH240_306K22,
genomic survey sequence.

CC479360

CC479360.1 GI:31756840

GSS.

Bos taurus (cow)

Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovidae; Bovinae; Bos.

1 (bases 1 to 949)

Holt, R., Stott, J., Yang, G., Barber, S., Smalls, D., Prabhu, A.-L.,

Tsai, M., Cloutier, A., Lee, D., Gern, N., Olson, T., Mayo, M.,

Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,

Matheson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,

Schein, J., Marra, M., de Jong, P., McWilliam, S., Barris, W.,

Dairymple, B.P. and Tellam, R.

Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398

Unpublished (2003)

Other GSSs: CH240_306K22.TARBAC13P2

Contact: Rob Holt

Sequencing

The British Columbia Cancer Agency Genome Science Centre

600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4B6

Tel: 604-877-6085

Fax: 604-877-6276

Email: rholt@bccsc.ca

Clones are derived from the bovine BAC library CHORI-240

TITLE
JOURNAL
COMMENT

REFERENCE
AUTHORS

ORGANISM

KEYWORDS

VERSION

DEFINITION

LOCUS

RESULT 3

```

Db      659 CAACAAGGTTTCATCTAGTCAAGGCTATGGTTTTTCCTGTCATGTATGGATGTGAGA 718
QY      2384 GTTGGACTGTGAAGAAAGCTGAGCACTGAAGAAATTGATGCTTTTGAACCTGTGGTTGGA 2443
Db      719 TTGGAAGTGTGAAGAGGCTGAGCAGCCGAGAAATGATGCTTTTGAACCTGTGGTTGGA 778
QY      2444 GAAGACTCTTGAAGAGTCCCTTGGACTGCAAGAGAGATCCAAACAGTCCATTCCTGAAGGAGA 2503
Db      779 GAAGACTCTTGAAGAGTCCCTTGGACTGCAAGAGAGATCCAAACAGTCCATTCCTGAAGGAGA 838
QY      2504 TCAGCCCTGGGATTTCTTTGGAGAAATGATGCTTAAGCTGAAGCTCCAGTACTTTGGCC 2563
Db      839 TCAGCCCTGGGATTTCTTTGGAGAAATGATGCTTAAGCTGAAGCTCCAGTACTTTGGCC 898
QY      2564 ACCTGATCAGAGAGCTGACTCACTCGGAAAGACCCCTGATGCTGGGAGGA 2614
Db      899 ACCTCATGCAAGAGTGTGACTCATTTGGAAAGAGACTCTGATGCTGGGAGGA 949

RESULT 4
CC508172/c
LOCUS   CH240_350K18.T7 CHORI-240 Bos taurus genomic clone CH240_350K18,
DEFINITION genomic survey sequence.
ACCESSION CC508172
VERSION   CC508172.1 GI:31826465
KEYWORDS  GSS.
SOURCE    Bos taurus (cow)
ORGANISM  Bos taurus
REFERENCE
AUTHORS   Holt,R., Stott,J., Yang,G., Barber,S., Smailus,D., Prabhu,A.-L.,
          Tsai,M., Cloutier,A., Lee,D., Girn,N., Olson,T., Mayo,M.,
          Butterfield,X., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R.,
          Mathewson,C., Wye,N., Masson,A., Brown-John,M., Jones,S.,
          Schein,J., Marra,M., de Jong,P., McWilliam,S., Barris,W.,
          Dairymple,B.P. and Tellam,R.
          Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
          Unpublished (2003)
          Other GSSs: CH240_350K18.TARBAC13P2
          Contact: Rob Holt
          Sequencing
          The British Columbia Cancer Agency Genome Science Centre
          600 W.10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
          Tel: 604-877-6085
          Fax: 604-877-6276
          Email: rholt@bcgsc.ca
          Clones are derived from the bovine BAC library CHORI-240
          (http://www.chori.org/bacpac/bovine240.htm). For BAC library
          availability, please contact Pieter de Jong (pjejong@mail.cho.org).
          Clones may be purchased from BACPAC Resources
          (http://www.chori.org/bacpac/ordering_information.htm). This work
          was undertaken as part of the International Bovine BAC Mapping
          Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
          British Columbia Genome Sciences Centre, Canada.
          Plate: 350 row: K column: 18
          Seq primer: T7
          Class: BAC ends.
FEATURES
            Location/Qualifiers
            1..927
               /organism="Bos taurus"
               /mol_type="genomic DNA"
               /strain="breed: Hereford"
               /db_xref="taxon:9913"
               /clone="CH240_350K18"
               /sex="Male"
               /cell_type="Blood"
               /clone_lib="CHORI-240"
               /note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
               Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC

```

Library (Male) produced by Pieter de Jong

Query Match 23.0%; Score 821.9; DB 1; Length 927;
 Best Local Similarity 93.8%; Pred. No. 0.11; Indels 2;
 Matches 876; Conservative 0; Mismatches 51; Gaps 7;

```

QY      408 GTCAAGTGTGCCCCATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGAATGA 467
Db      927 GTCAAGTGTGCCCCATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGAATGA 868
QY      468 AGAGATGGAGCCCAACCAAGAAAGATACCCAGCTGTGGATGTGACTGGTGAATAGCA 527
Db      867 AGGAATGGAGCCCAAGCAAAACAATACCCAGCTGTGGATGTGACTGGTGAATAGCA 808
QY      528 GTTCCGATCTGTAAAGAGCAATATTCATAGAGAACTGGAATGTGAGTCCATGAATCA 587
Db      807 GTTCCGACGCTGTAAAGAGCAATATTCATAGAGAACTGGAATGTGAGTCCATGAATCA 748
QY      588 AGSCAAATTTGGAAGTGTGTCAAACAAGAGATGGCAAGAGTGAATGTCAACATTTCTAGGAAT 647
Db      747 AGSCAAATTTGGAAGTGTGTCCAACAAGAGATGGCAAGAGTGAATGTGAGTCCATGAATCA 688
QY      648 CAGCGAACTAAAATGGAGTGGAAATGGTGAATTTAACTCAGATGACCATTTATATCTACTA 707
Db      687 CAGCGAACTAAAATGGAGTGGAAATGGTGAATTTAACTCAGATGACCATTTATATCTACTA 628
QY      708 CTGCGGGCAGGAATCCCTCAGAGAAATGGAGTAGCCATCATGTCACAAAGAGTCCG 767
Db      627 CTGCGGGCAGGAATCCCTCAGAGAAATGGAGTAGCCATCATGTCACAAAGAGTCCG 568
QY      768 AAATCAGTACTTGGATCGAGTCTCAAAAACAGCAGAGATGATCTCTGTTGTTCCAGG 827
Db      567 AAATCAGTACTTGGATCGAATCTCAAAAACAGCAGAGATGATCTCTGTTGTTCCAGG 508
QY      828 CAAACCATTTCAATATCACAGTAATCCAAGTCTCTATCCCCAACCCAGTAATGTCTGAAGAAGC 887
Db      507 CAAACCATTTCAATATCACAGTAATCCAAGTCTCTATGCCCAACCAGTAAGGCTGAAGAAGC 448
QY      888 TGAAGTTGACGGTCTCTATGAAGACTACAGACCTTTTAGACTTAACACCCCAAAAAGA 947
Db      447 -----TGAACGGTTCTATAGAGACCTTACAGGGCCCTTTTAGAAGTAAACCCCAAAAAGA 394
QY      948 TGTCTCTTCTCATTTATAGGGGACTGGAATGCAAAAGTAGGAAGCAAGAAACAACCTGGAGT 1007
Db      393 TGTCTCTTTTAAATTTAGGGGACT-GAGTGCATAAGTAGGAATCAAGAAACACCTGGAGT 335
QY      1008 AACAGGCAAAATTTGGCTTTGGAATACGGAATGAAGCAGGGCAAGAGCTAATAGAGTTTTG 1067
Db      334 AACAGGCAAAATTTGGCTTTGGAATACGGAATGAAGCAGGGCAAGAGCTAATAGAGTTTTG 275
QY      1068 CCAAGAAATGCACTGGTTCATAGCAAAACACCCCTCTTCCAAACAACAGAGAGAGACTCTTA 1127
Db      274 CCAAGAAATGCACTGGTTCATACAAATACCCCTCTTCCAAACAACAGAGAGAGACTCTTA 215
QY      1128 CACATGGACATCACAGATGGTTCACACCCGAAATCAGATTGATTATATTTTTCAGAGCA 1187
Db      214 CACATGGACATCACAGATGGTTCACACCCGAAATCAGATTGATTATATTTTTCAGAGCA 155
QY      1188 AAGATGGAGAGCTCTATACAGTCAAGCAAAACAGACCCAGGAGCTTACTGTGGCTCAGA 1247
Db      154 AAGATGGAGAGCTCTATACAGTCAAGCAAAACAGACCCAGGAGCTTACTGTGGCTCAGA 95
QY      1248 TCATGAACCTCTTATTTGCCAAATTCAGACTTTAAATTTGAAGAAAGTAGGGAAAACCACTAG 1307
Db      94 TCATGAACCTCTTATTTGCCAAATTCAGACTTTAAATTTGAAGAAAGTAGGGAAAACCACTAG 35
QY      1308 ATCACTCAGGTAAAGACCTTAATTCNAATCCCTTAT 1341
Db      34 GCCATTTCAGGTATGACCTTAATTCNAATCCCTTAT 1

```

RESULT 5
 CC545978

LOCUS CC544978 894 bp DNA linear GSS 17-JUN-2003
DEFINITION CH240_427A14.TARBAC13P2 CHORI-240 Bos taurus genomic clone
CH240_427A14, genomic survey sequence.

ACCESSION CC544978
VERSION CC544978.1 GI:31863262
KEYWORDS GSS.
SOURCE Bos taurus (cow)

ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 894)
AUTHORS Holt, R., Stott, J., Yang, G., Barber, S., Smailus, D., Prabhu, A.-L.,
Tsai, M., Cloutier, A., Lee, D., Girn, N., Olson, T., Mayo, M.,
Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
Schein, J., Marra, M., de Jong, P., Keefe, J.W. and Kappes, S.M.

TITLE Bovine BAC End Sequences from Library CHORI-240, PLATES 399 to 478
JOURNAL Unpublished (2003)
COMMENT Other GSSs: CH240_427A14.T7
Contact: Rob Holt

Sequencing
The British Columbia Cancer Agency Genome Science Centre
600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
Tel: 604-877-6085
Fax: 604-877-6276
Email: rholt@bcgsc.ca

Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering/information.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
British Columbia Genome Sciences Centre, Canada.

Plate: 427 Row: A Column: 14
Seq primer: SP6
Class: BAC ends.

Location/Qualifiers
1..894
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="Breed: Hereford"
/db_xref="taxon:9913"
/clones="CH240_427A14"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
Hereford bull LI Domino 99375; CHORI-240 Bovine BAC
Library (Male) produced by Pieter de Jong"

Query Match 22.9%; Score 818.2; DB 1; Length 894;
Best Local Similarity 96.0%; Pred. No. 0.12;
Matches 849; Conservative 0; Mismatches 33; Indels 2; Gaps 1;

QY 298 GTGTCCTCACTGAGAGGAAT--GCAACCACTTCAGTATCTTGGCTTGAGAACCCCA 355
DB 11 GTGTCCTCACTGAGAGGAATGCAACCACTTCAGTATCTTGGCTTGAGAACCCCA 70

QY 356 TGAACAGATGAAAAGCAAAATGATAGTACTGAAAGAGCACTCCCAAGTCACTAG 415
DB 71 TGAACAGATGAAAAGCAAAATGATAGTACTGAAAGAGCACTCCCAAGTCACTAG 130

QY 416 GTGCCCATATGCTACTGAGATCAGTGGAGAAATACTCCAGAAAGATGAAGATGG 475
DB 131 GTGCCCATATGCTACTGAGATCAGTGGAGAAATACTCCAGAAAGATGAAGATGG 190

QY 476 AGCCAAAGCAAAAGAAATACCCAGCTGGGATGCTGCTGGTATAGCATAGCAAGTCCGAT 535
DB 151 AGCCAAAGCAAAAGAAATACCCAGCTGGGATGCTGCTGGTATAGCATAGCAAGTCCGAT 250

QY 536 GCTGTAAGAGCAATATTGTCATAGAACCTCGAATGTCAGTCCATGAATCAAGGCAAT 595

DB 251 GCTGTAAGAGCAATATTGTCATAGAACCTCGAATGTCAGTCCATGAATCAAGGCAAT 310

QY 596 TGAAGTGTCTAAACAAGAGATGGCAAGAGTGAATGTCAACATTTAGGAATCAGGCAAC 655
DB 311 TGAAGTGTCTAAACAAGAGATGGCAAGAGTGAATGTCAACATTTAGGAATCAGGCAAC 370

QY 656 TAAAATGGACTGGAATGGGTGAATTTAACTCAGATGACCATATATCTACTACTCGGGC 715
DB 371 TGAATGGACTGGAATGGGTGAATTTAACTCAGATGACCATATATCTACTACTCGGGC 430

QY 716 AGGAATCCCTCAGAAGAAATGGAGTAGCCATCATGTCAACAAAAGAGTCCGAAATGCGAG 775
DB 431 AGGAATCCCTCAGAAGAAATGGAGTAGCCATCATGTCAACAAAAGAGTCCGAAATGCGAG 490

QY 776 TACTTGGATGCGAGTCTCAAAAACGACAGATGATCTCTGTTTGTTCCTCAAGGCAACCAT 835
DB 491 TACTTGGATGCGAGTCTCAAAAACGACAGATGATCTCTGTTTGTTCCTCAAGGCAACCAT 550

QY 836 TCAATATCACAGTAATCCAAAGTCTATGCCCCCAACAGTAATGCTGAAGAGCTGAAGTTG 895
DB 551 TCAATATCACAGTAATCCAAAGTCTATGCCCCCAACAGTAATGCTGAAGAGCTGAAGTTG 610

QY 896 AAGGTCCTATGAGACCTTACAGACCTTTAGAACCTTAACCCCAAAAAGAGTCTCTTC 955
DB 611 AAGGTCCTATTAAGACCTTACAGACCTTTAGAACCTTAACCCCAAAAAGAGTCTCTTC 670

QY 956 TCATTATAGGGGACTGGAATGCAAAAAGTAGGAAGCAAAAGAACACCTCGAGTAACAGGCA 1015
DB 671 TCATTATAGGGGACTGGAATGCAAAAAGTAGGAAGCAAAAGAACACCTCGAGTAACAGGCA 730

QY 1016 AATTGGCTTGGAAATACGGAATGAAGAGGGCAAGAGTAATAGAGTTTGGCAAGAA 1075
DB 731 AATTGGCTTGGAAATACGGAATGAAGAGGGCAAGAGTAATAGAGTTTGGCAAGAA 790

QY 1076 ATGCACTGGTCTAGCAAAACACCTCTTCCAAACACAGAGAGACCTCTACACATGA 1135
DB 791 ATGCACTGGTCTAGCAAAACACCTCTTCCAAACACAGAGAGACCTCTACACATGA 850

QY 1136 CATCACAGATGCTCAACACCCGAAATCAGATTGATTATTTCT 1179
DB 851 CATCACAGATGCTCAACACCCGAAATCAGATTGATTATTTCT 894

RESULT 6
CC577078/c 919 bp DNA linear GSS 18-JUN-2003
LOCUS CH240_455012.TARBAC13P2 CHORI-240 Bos taurus genomic clone
DEFINITION CH240_455012, genomic survey sequence.

ACCESSION CC577078
VERSION CC577078.1 GI:31920393
KEYWORDS GSS.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 919)
AUTHORS Holt, R., Stott, J., Yang, G., Barber, S., Smailus, D., Prabhu, A.-L.,
Tsai, M., Cloutier, A., Lee, D., Girn, N., Olson, T., Mayo, M.,
Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
Schein, J., Marra, M., de Jong, P., Keefe, J.W. and Kappes, S.M.

TITLE Bovine BAC End Sequences from Library CHORI-240, PLATES 399 to 478
JOURNAL Unpublished (2003)
COMMENT Other GSSs: CH240_455012.T7
Contact: Rob Holt

Sequencing
The British Columbia Cancer Agency Genome Science Centre
600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
Tel: 604-877-6085
Fax: 604-877-6276
Email: rholt@bcgsc.ca

Query Match 22.6%; Score 806.8; DB 1; Length 882;
Best Local Similarity 94.7%; Pred. No. 0.14; Indels 0; Gaps 0;
Matches 835; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

CC586882
VERSION
KEYWORDS
SOURCE
ORGANISM

CC586882.1 GI:31939990
GSS
Bos taurus (cow)
Bos taurus
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 866)
Holt, R., Cloutier, J., Yang, G., Barber, S., Smailus, D., Prabhu, A., Tsai, M., Scott, J., Lee, D., Girn, N., Olson, T., Mayo, M., Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R., Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S., Schein, J., Marra, M., de Jong, P., McWilliam, S., Barris, W., Dailympie, B. P. and Tellam, R.
Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
Unpublished (2003)
Other GSSs: CH240_384H21.TARBAC13P2
Contact: Rob Holt
Sequencing
The British Columbia Cancer Agency Genome Science Centre
600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
Tel: 604-877-6085
Fax: 604-877-6276
Email: rholt@bcgsc.ca
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/orderinginformation.html). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
British Columbia Genome Sciences Centre, Canada.
Plate: 384 row: H column: 21
Seq primer: T7
Class: BAC ends.

FEATURES
source
1..866
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="breed: Hereford"
/db_xref="taxon:9913"
/clone="CH240_384H21"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/notes="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
Hereford bull Li Domino 99375; CHORI-240 Bovine BAC
library (Male) produced by Pieter de Jong"

Query Match 22.4%; Score 801.7; DB 1; Length 866;
Best Local Similarity 96.1%; Pred. No. 0.15; Indels 1; Gaps 1;
Matches 832; Conservative 0; Mismatches 33; Indels 1; Gaps 1;

Qy 510 GACTGTGATATAAGCAAGTCCGATGCTGTAAAGAGCAATATTCATAGGAACCTGGAA 569
Db 866 GACTGTGATAGAAGCAAGTCCGATGCTGTAAAGAGCAATATTCATAGGAACCTGAA 807
Qy 570 TGTCAAGTCCATCAATCAAGGCAATGGAAGTGTCAACAGAGATGCCAGAGTGAA 629
Db 806 TGTCAAGTCCATCAATCAAGGCAATGGAAGTGTCAACAGAGATGCCAGAGTGAA 747
Qy 630 TGTCAACATTCAGGAATCAGCAACTAAAAATGGACTGGAATGGGTGAATTTAACTCAGA 689
Db 746 TGTCAACATTCAGGAATCAGCAACTAAAAATGGACTGGAATGGGTGAATTTAACTCAGA 687
Qy 690 TGACCATTTATCTACTCTGCGGAGGAATCCCTCAGAAGAAATGGAGTAGCCATCAT 749
Db 686 TGACCATTTATCTACTCTGCGGAGGAATCCCTTAGAAGAAATGGAGTAGCCATCAT 627
Qy 750 GGTCAACAAAGAGTCCGAAATGCAGTACTTGGATCGAGTCTCAAAAACGACAGAAATGAT 809
Db 626 GGTCAACAAAGAGTCCGAAATGCAGTACTTGGATCGAGTCTCAAAAATGCAGAAATGAT 567

RESULT 8

CC586882/c

LOCUS

DEFINITION

CC586882
CH240_384H21.T7 CHORI-240 Bos taurus genomic clone CH240_384H21,
genomic survey sequence.

QY 810 CTCTGTTTGTTCCTCAAGCGCAACCAATTCATATACAGTATCAAGTCTATGCCCAAC 869
 Db 566 CTCTGTTTGTTCCTCAAGCGCAACCAATTCATATACAGTATCAAGTCTATGCCCAAC 507
 QY 870 CAGTAATGCTGAAGAGCTGAAGTTGAACGGTCTCTATGAAGACCTTACAGACCTTTTGA 929
 Db 506 CAGTAACACTGAAGAGCTGAAGTTGAACGGTCTCTATGAAGACCTTACAGACCTTTTGA 447
 QY 930 ACTACACCCCAAAAGAGATGCTTCTATATATAGGGAAGTGGATGCAAAAGTAGGAG 989
 Db 446 ACTACACCCCAAAAGAGATGCTTCTATATATAGGGAAGTGGATGCAAAAGTAGGAG 388
 QY 990 CAAAGAAACACTGAGTAACAGGCAAAATTTGGCTTGGATACGGAATGAAGCAGGCA 1049
 Db 387 TCAAGAAATACCTGGAGTAACAGGCAAAATTTGGCTTGGATACGGAATGAAGCAGGCA 328
 QY 1050 AAGACTAATAGAGTTTTCGCAAGAAATGCACTGCTGTATAGCAACACCTCTTCCAACA 1109
 Db 327 AAGGCTAATAGAGTTTTCGCAAGAAATGCACTGCTGTATAGCAACACCTCTTCCAACA 268
 QY 1110 ACACAAGAGAGACTCTACATGACATCACCAGATGCTCAACACCGAATCAGATTGA 1169
 Db 267 ACACAAGAGAGACTCTACATGACATCACCAGATGCTCAACACCGAATCAGATTGA 208
 QY 1170 TTATATCTTTTCGCAAGAAATGCACTGCTGTATAGCAACACCTCTTCCAACA 1229
 Db 207 TTATATCTTTTCGCAAGAAATGCACTGCTGTATAGCAACACCTCTTCCAACA 148
 QY 1230 AGCTTACTGCTCAGATCATGACTCTTATGCTGCAATTCAGACTTAATTTGAAGA 1289
 Db 147 AGCTTACTGCTCAGATCATGACTCTTATGCTGCAATTCAGACTTAATTTGAAGA 88
 QY 1290 AGTAGGGAACCACTAGATCACTCAGGTAAGAGCTTAATCCAATCCCTTATGATTATAC 1349
 Db 87 AGTAGGGAACCACTAGATCACTCAGGTAAGAGCTTAATCCAATCCCTTATGATTATAC 28
 QY 1350 AGTGAAGTGAATAGATTAAAG 1375
 Db 27 AGTGAAGTGAATAGATTAAAG 2
 RESULT 9
 CC520949
 LOCUS CH240_368M20.TARBA13P2 CHORI-240 Bos taurus genomic clone
 DEFINITION CH240_368M20, genomic survey sequence.
 ACCESSION CC520949
 VERSION CC520949.1 GI:31839237
 KEYWORDS GSS
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 REFERENCE 1 (bases 1 to 887)
 AUTHORS Holt, R., Stott, J., Yang, G., Barber, S., Smalhus, D., Prabhu, A.-L.,
 Tsai, M., Cloutier, A., Lee, D., Girm, N., Olson, T., Mayo, M.,
 Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
 Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
 Schein, J., Marra, M., de Jong, P., McWilliam, S., Barris, W.,
 Dalrymple, B.P., and Tellam, R.
 TITLE Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
 JOURNAL Unpublished (2003)
 COMMENT Other GSSs: CH240_368M20.77
 Contact: Rob Holt
 Sequencing The British Columbia Cancer Agency Science Centre
 600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4S6
 Tel: 604-877-6085
 Fax: 604-877-6276
 Email: rholt@bccsc.ca
 Clones are derived from the bovine BAC library CHORI-240

(http://www.chori.org/bacpac/bovine240.htm). For BAC library
 availability, please contact Pieter de Jong (pdejong@mail.cho.org).
 Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/ordering/information.htm). This work
 was undertaken as part of the International Bovine BAC Mapping
 Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
 British Columbia Genome Sciences Centre, Canada.
 Plate: 368 row: M column: 20
 Seq primer: SP6
 Class: BAC ends.

FEATURES source

Location/Qualifiers
 1..887
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /strain="breed: Hereford"
 /db_xref="taxon:9913"
 /clone="CH240_368M20"
 /sex="Male"
 /cell_type="Blood"
 /clone_lib="CHORI-240"
 /note="Vector: pTARBA1.3; Site 1: MboI; Site 2: MboI;
 Hereford bull Ll Domino 99375; CHORI-240 Bovine BAC
 library (Male) produced by Pieter de Jong"

Query Match 22.4%; Score 800.3; DB 1; Length 887;
 Best Local Similarity 94.6%; Pred. No. 0.14;
 Matches 839; Conservative 0; Mismatches 47; Indels 1; Gaps 1;
 QY 289 TGACAGAATGTGTCACCTGGAGAGAGGAAT-GCAAAACACCTTCAGTATTCCTGCTTGA 347
 Db 1 TGACAGAATGTGTCACCTGGAGAGAGTAATGGCAACACCTTCAGTATTCCTGCTTGA 60
 QY 348 GAACCCCATGAACAGTATGAAAGGCAAAATGATAGGATCTGAAAGAGGAACCTCCCAAG 407
 Db 61 GAACCCCATGAACAGTATGAAAGGCAAAATGATAGGATCTGAAAGAGGAACCTCCCAAG 120
 QY 408 GTCACTAGTGTGCCCCATATGCTCTGGAGATCAGTGGGAATACTCCAGAAAGAATGA 467
 Db 121 GTCACTAGTGTGCCCCATATGCTCTGGAGATCAGTGGGAATACTCCAGAAAGAATGA 180
 QY 468 AGAGATGGAGCCAAAGCAAAAGAGATACCCAGCTGTGGATGTGACTGGTGTATATAGCAA 527
 Db 181 AGGACGGAACCAAGCAACACACACCCAGCTGTGGATGTGACTGGTGTATAGCAA 240
 QY 528 GGTCCGATCTGTAAAGCAATTTGCATAGGAACCTGGAAATGTCAGTCCATGAATCA 587
 Db 241 GGTCCGATCTGTAAAGCAATTTGCATAGGAACCTGGAAATGTCAGTCCATGAATCA 300
 QY 588 AGGCAAAATTTGGAAGTGGTCAAAAGAGATGGCAAGATGGAATGTCAACATCTTAGGAAT 647
 Db 301 AGGCAAAATTTGGAAGTGGTCAAAAGAGATGGCAAGATGGAATGTCAACATCTTAGGAAT 360
 QY 648 CAGCGAACTTAAATGGACTGGAAATGGGTGAATTTAACTCAGATGACCATTTATCTACTA 707
 Db 361 CAGCGAACTTAAATGGACTGGAAATGGGTGAATTTAACTCAGATGACCATTTATCTACTA 420
 QY 708 CTGCGGCGAGGAATCCCTCAGAAAGAAATGGAGTAGCCATCATGGTCAACAAAGAGTCCG 767
 Db 421 CTGCGGCGAGGAATCCCTCAGAAAGAAATGGAGTAGCCATCATGGTCAACAAAGAGTCCG 480
 QY 768 AAATGCAGTACTTGGATGCAGTCTCAAAAGCAAGAGATGATCTCTGTGTTCCTCAAGG 827
 Db 481 AAATGCAGTACTTGGATGCAGTCTCAAAAGCAAGAGATGATCTCTGTGTTCCTCAAGG 540
 QY 828 CAAACCAATTCATATCAGAGTAATCCAGTCTATGCCCAACCATGATGCTGAGAGGC 887
 Db 541 CAAACCAATTCATATCAGAGTAATCCAGTCTATGCCCAACCATGATGCTGAGAGGC 600
 QY 888 TGAAGTTGAACCGTCTCTATGAGAGCTTCAAGACCTTTTAGAAGCTTAACACCAAAAAAGA 947
 Db 601 TGAAGTTGAAGTCTTATGAAGATCTATAAGATCTTTTAGAAGCTTAACACCAAAAAAGA 660
 QY 948 TGTCTCTTCTATTATAGGGGACTGGAATGCAAAAGTAGGAAGCAAGAAACACCTGGAGT 1007

Db 661 TGTCTTTTATATAGTGTGATGGAATGGAAGTAGGAGTCAAGAAACATCTGGAGT 720
 Qy 1008 AACAGGCAAAATTTGGCTTGAATACGAATGAAGAGGGCAAGAGCTAATAGAGTTTG 1067
 Db 721 AACAGGCAAAATTTGGCTTGAATACGAATGAAGAGGGCAAGAGCTAATAGAGTTTG 780
 Qy 1068 CCAAGAAATCACTGGTCAATAGCAAAACACCTCTTCCAAACACAGAGAGAGACTTA 1127
 Db 781 CCAAGAAATCACTGGTCAATAGCAAAACACCTCTTCCAAACACAGAGAGAGACTTA 840
 Qy 1128 CACATGGACATCAACAGAGTGTCAACACCGAAATCAGATTGATTATA 1174
 Db 841 TACATGGACATCAACAGAGTGTCAACACCGAAATCAGATTGATTATA 887

RESULT 10
 CC520792/c 951 bp DNA linear GSS 17-JUN-2003
 LOCUS CH240_368C2.T7 CHORI-240 Bos taurus genomic clone CH240_368C2,
 DEFINITION genomic survey sequence.

ACCESSION CC520792
 VERSION CC520792.1 GI:31839080
 KEYWORDS GSS.
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 951)

REFERENCE
 AUTHORS Holt, R., Stott, J., Yang, G., Barber, S., Smalls, D., Prabhu, A.-L.,
 Tsai, M., Cloutier, A., Lee, D., Girn, N., Olson, T., Mayo, M.,
 Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
 Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
 Schein, J., Marra, M., de Jong, P., McWilliam, S., Barris, W.,
 Dalrymple, B. P. and Tellam, R.
 Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
 Unpublished (2003)
 JOURNAL Other GSSs: CH240_368C2.TARBAC13P2
 COMMENT Contact: Rob Holt

Sequencing The British Columbia Cancer Agency Genome Science Centre
 600 W. 10th Ave Vancouver, British Columbia, Canada V5Z 4B6
 Tel: 604-877-6085
 Fax: 604-877-6276
 Email: rholt@bccsc.ca
 Clones are derived from the bovine BAC library CHORI-240
 (<http://www.chori.org/bacpac/bovine240.htm>). For BAC library
 availability, please contact Pieter de Jong (pdejong@mail.cho.org).
 Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/ordering_information.htm). This work
 was undertaken as part of the International Bovine BAC Mapping
 Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
 British Columbia Genome Sciences Centre, Canada.
 Plate: 368 row: C column: 2
 Seq primer: 77
 Class: BAC ends.
 Location/Qualifiers
 1..951
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /strain="bred; Hereford"
 /db_xref="taxon:9913"
 /clone="CH240_368C2"
 /sex="Male"
 /cell_type="Bicoid"
 /clone_lib="CHORI-240"
 /note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
 Hereford bull Li Domino 99375; CHORI-240 Bovine BAC
 library (Male) produced by Pieter de Jong"

FEATURES

source

Query Match 22.4%; Score 800.1; DB 1; Length 951;
 Best Local Similarity 93.5%; Pred. No. 0.13;

Matches 886; Conservative 0; Mismatches 55; Indels 7; Gaps 5;
 Qy 1811 TTGTATGAGTGTGAGAGCAACAGTTAGAACTGGACATGGAACACACAGACTGGTTCCAA 1870
 Db 951 TTGTATGAGTGTGAGAGCAACAGTTAGAACTGGACAT-GAACACAGACT-GTTCCAA 894
 Qy 1871 GTAGGAAAAGGAGTATGTCAAGCTGTATATGTACCCGGCTTGTAACTTCTATGCA 1930
 Db 893 AGAGGAAAAGGAGTATGTCAAGCTGTATATGTACCCGTCTTTTAACTTATATGCA 834
 Qy 1931 GAG-ACATCATGAGAAACGCTGGCTGGAAGAGCAAAAGCTGGAATCAAGATTGCGGG 1989
 Db 833 GAGTATCATGAGAAACGCTGGCTGGAAGAGCAAAAGCTGGAATCAAGATTGCGGG 774
 Qy 1990 AGAATAGCAATAAACCTCAGATATGAGATATACCAACCTTATGSCAGAAAGTGAAGAG 2049
 Db 773 AGAATATCAATTAACCTCAGATATGAGATATACCAACCTTATGSCAGAAAGTGAAGAG 714
 Qy 2050 GAACATAAGGCTCTTTGATGAAGTGAAGAGGAGAGTGAAGAAAGTTGCTTAAAGCTC 2109
 Db 713 GAACATAAGGCTCTTTGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 654
 Qy 2110 AACATTGAGAAACGAGATCATGGCATCTGGTCCCATCATTCTGGAAGAAATAGATGG 2169
 Db 653 AACATTGAGAAACGAGATCATGGCATCTGGTCCCATCATTCTGGAAGAAATAGATGG 594
 Qy 2170 GAAACAGTGGAAACAGTGTGAGACTTTATTTTGGGGGGCTCCAAATCACTGAGATG 2229
 Db 593 GAAACAGTGGAAACAGTGTGAGACTTTATTTT-CTGGGCTCCAAATCACTGAGATG 536
 Qy 2230 GTGACTGAGCCATGAATTAAGACATCTACTCTTGGAGAAAGTTA-ACCAAC 2287
 Db 535 GTGACTGAGCCATGAATTAAGACGCTTACTCTTGGAGAAAGTTATGACCAAC 476
 Qy 2288 TAGATAGCATATTGAAAAGCAGAGACATTACCTTGCCAAAGCCCATCTAGTCAAGG 2347
 Db 475 TAGATAGCATATTCAAGACAGACATTACTTTGCCAACAAGGTTCTGCTAGTCAAGG 416
 Qy 2348 CTATGTTTTTCCAGTGGTCAATGTATGATGTGAGAGTTGGAGCTGTGAAGAAAGCTGAGC 2407
 Db 415 CTATGTTTTTCCAGTGGTCAATGTATGATGTGAGAGTTGGAGCTGTGAAGAAAGCTGAGC 356
 Qy 2408 ACTGAGAAATGATGCTTTTGAACCTGGTGTGGAGAGAGACTCTTGGAGAGTCCCTTGA 2467
 Db 355 GCCGAAGAAATGATGCTTTTGAACCTGGTGTGGAGAGAGACTCTTGGAGAGTCCCTTGA 296
 Qy 2468 CTGCAAGGAGATCCAAACCACTCCATTTCTGAAGAGATCAGCCCTGGGATTTCTTTGGAAG 2527
 Db 295 CTGCAAGGAGATCCAAACCACTCCATTTCTGAAGAGATCAGCCCTGGGATTTCTTTGGAAG 236
 Qy 2528 GAATGATGCTAAAGCTGAAACCTCCAGTACTTTGGCCACCTGATCAGAAAGCTGACTCAG 2587
 Db 235 GAATGATGCTAAAGCTGAAACCTCCAGTACTTTGGCCACCTGATCAGAAAGCTGACTCAG 176
 Qy 2588 TGGAAAAGACCCCTGATGCTGGGAGGATTTGGGGCAGGAGGAGAGCCGACACAGAGGA 2647
 Db 175 TGGAAAAGACCTGATGCTGGGAGGATTTGGGGCAGGAGGAGAGCCGACACAGAGGA 116
 Qy 2648 TGAGATGGCTGGATGCGCATCACTCAATCGATCGAGCTGAGTCTGGGTGAACCTCTGGAGT 2707
 Db 115 TGAGATGGCTGGATGCGCATCTGCTGATGAGCGTGTGAGTCTCCCTGAACTCCCGGAGT 56
 Qy 2708 TGGTGTGAGCAGGAGGAGGCTGTCTCGCGCGGATTCATCGGGTCAAA 2755
 Db 55 TGGTGTGAGCAGGAGGAGGCTGTCTGCTGTGATTATGAGGTGCAAA 8

RESULT 11
 CC548437
 LOCUS CH240_432010.T7 CHORI-240 Bos taurus genomic clone CH240_432010,
 DEFINITION genomic survey sequence.
 ACCESSION CC548437

CC548437 836 bp DNA linear GSS 17-JUN-2003
 CH240_432010.T7 CHORI-240 Bos taurus genomic clone CH240_432010,
 genomic survey sequence.
 CC548437

```

VERSION CC548437.1 GI:31866721
KEYWORDS GSS.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus

REFERENCE
AUTHORS Holt,R., Cloutier,J., Yang,G., Barber,S., Smalilus,D., Prabhu,A.-L.,
Tsai,M., Cloutier,A., Lee,D., Girn,N., Olson,T., Mayo,M.,
Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R.,
Mathewson,C., Wye,N., Masson,A., Brown-John,M., Jones,S.,
Schein,J., Marra,M., de Jong,P., Keele,J.W. and Kappes,S.M.
Bovine BAC End Sequences from Library CHORI-240, PLATES 399 to 478
Unpublished (2003)
Other GSSs: CH240_432010.TARBAC13P2
Contact: Rob Holt
Sequencing
The British Columbia Cancer Agency Genome Science Centre
600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4B6
Tel: 604-877-6085
Fax: 604-877-6276
Email: rholt@bcgsc.ca
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering/information.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
British Columbia Genome Science Centre, Canada.
Plate: 432 row: O column: 10
Seq primer: 17
Class: BAC ends.

FEATURES
source
1. 836
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="breed: Hereford"
/db_xref="taxon:9913"
/clone="CH240_432010"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
library (Male) produced by Pieter de Jong"

Query Match 22.3%; Score 798.2; DB 1; Length 836;
Best Local Similarity 97.2%; Pred. No. 0.16;
Matches 812; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 320 GCAACCACTTCAGTATTCCTTGGCTTGAGACCCCATGAAACAGTATGAAGGCAAAATG 379
DB 2 GCAACCACTTCAGTATTCCTTGGCTTGAGACCCCATGAAACAGTATGAAGGCAAAATG 61
QY 380 ATAGATATCTGAAGAGAACTCCCGAGTCAGTAGGTGGCCCATATGCTACTGGAGATC 439
DB 62 ATAGATATCTGAAGAGAACTCCCGAGTCAGTAGGTGGCCCATATGCTACTGGAGATC 121
QY 440 AGTGAGAAATATCTCCGAAGAATGAAGATGGAGCCCAAGCAAAAGAAATACCCAG 499
DB 122 AGTGAGAAATATCTCCGAAGAATGAAGATGGAGCCCAAGCAAAAGAAATACCCAG 181
QY 500 CTGTGGATGTGACTGGTGATATAAGCAAGGTCCTGATGCTGTAAGAGCAATATTGCATAG 559
DB 182 CTGTGGATGTGACTGGTGATATAAGCAAGGTCCTGATGCTGTAAGAGCAATATTGCATAG 241
QY 560 GAACCTGGAATGTCCAGGTCCATGAATCAAGGCAAAATGGAGTGGTCAAAACAGAGATGG 619
DB 242 GAACCTGGAATGTCCAGGTCCATGAATCAAGGCAAAATGGAGTGGTCAAAACAGAGATGG 301
QY 620 CAAGAGTGAATCTCAACATTCTAGGAATCAGCGAACTAAATGGACTGGATGGTGAAT 679

```

```

DB 302 CAAGAGTGAATGCGACATTTAGGAATCAGCGAACTGAAATGGACTGGAATGGTGAAT 361
QY 680 TTAACCTCAGATGACCATTTATATCTACTCTGCGGAGGAATCCCTCAGAGAAATGGAG 739
DB 362 TTAACCTCAGATGACCATTTATATCTACTCTGCGGAGGAATCCCTCAGAGAAATGGAG 421
QY 740 TAGCCATCATGTGTCAACAAAAGAGTCCGAAATCGAGTACTTGGATGCAATCTCAAAAACG 799
DB 422 TAGCCATCATGTGTCAACAAAAGAGTCCGAAATCGAGTACTTGGATGCAATCTCAAAAACG 481
QY 800 ACAGATGATCTCTGTTTCTTCCAGGCAAAACCATTCGAATATCAGATTAATCCAGTCT 859
DB 482 ACAGATGATCTCTGTTTCTTCCAGGCAAAACCATTCGAATATCAGATTAATCCAGTCT 541
QY 860 ATGCCCCCAACCAAGTATGCTGAAGAGCTGAAGTGTGAACGGTCTTATGAAGACCTACAAG 919
DB 542 ATGCCCCCAACCAAGTATGCTGAAGAGCTGAAGTGTGAACGGTCTTATGAAGACCTACAAG 601
QY 920 ACCTTTTAGAATCAACACCCAAAAGAGTGTCTTCTCATTTAGGGGAGTGGATGCA 979
DB 602 ACCTTTTAGAATCAACACCCAAAAGAGTGTCTTCTCATTTAGGGGAGTGGATGCA 661
QY 980 AAGTAGGAAGCAAGAAACACCTGGAGTAAACAGGAAATTTGGCCCTTGGAAATACGGAATG 1039
DB 662 AAGTAGGAAGTCAAGAAACACCTGGAGTAAACAGGCAATTTGGCCCTTGGAAATACGGAATG 721
QY 1040 AAGCAGGGCAAGACATATAGAGTTTTTGGCAAGAAATGCACTGTGTATAGCAACACCC 1099
DB 722 AAGCAGGGCAAGACATATAGAGTTTTTGGCAAGAAATGCACTGTGTATAGCAACACCC 781
QY 1100 TCTTCCAAACACACAGAGAGACCTCTACACATGGACATCACCAGATGGTCAACA 1154
DB 782 TCTTCCAAACACACAGAGAGACCTCTATACATGGACATCACCAGATGGTCAACA 836

CC594700 938 bp DNA linear GSS 18-JUN-2003
CH240_396F13.77 CHORI-240 Bos taurus genomic clone CH240_396F13,
genomic survey sequence.
CC594700
CC594700.1 GI:31955684
GSS.
Bos taurus (cow)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
Holt,R., Cloutier,J., Yang,G., Barber,S., Smalilus,D., Prabhu,A.-L.,
Tsai,M., Cloutier,A., Lee,D., Girn,N., Olson,T., Mayo,M.,
Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R.,
Mathewson,C., Wye,N., Masson,A., Brown-John,M., Jones,S.,
Schein,J., Marra,M., de Jong,P., McWilliam,S., Barris,W.,
Dalrymple,B.P. and Tellam,R.
Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
Unpublished (2003)
Other GSSs: CH240_396F13.TARBAC13P2
Contact: Rob Holt
Sequencing
The British Columbia Cancer Agency Genome Science Centre
600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4B6
Tel: 604-877-6085
Fax: 604-877-6276
Email: rholt@bcgsc.ca
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering/information.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the

```

British Columbia Genome Sciences Centre, Canada.
 Plate: 396 row: F column: 13
 Seq primer: T7
 Class: BAC ends.

FEATURES

Location/Qualifiers
 1. .938
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /strain="breed: Hereford"
 /db_xref="taxon:9913"
 /clone="CH240_396F13"
 /sex="Male"
 /cell_type="Blood"
 /clone_lib="CHORI-240"
 /note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
 Hereford bull Li Domino 99375; CHORI-240 Bovine BAC
 library (Male) produced by Pieter de Jong"

Query Match 22.3%; Score 798.2; DB 1; Length 938;
 Best Local Similarity 93.1%; Pred. No. 0.14; Mismatches 59; Indels 6; Gaps 4;
 Matches 876; Conservative 0;

290 GACAGATGTGCTCCACTGGAGAGGAA-TGCAACCACTTCAGTATTCTTGCTTGAG 348
 |||||
 938 GACACATGTGCTCCACTGGAGAGGAA-TGCAACCACTTCAGTATTCTTGCTTGAG 879
 |||||
 349 AACCCCATGAACAGTATGAAAAGGCAAAATGATAGGATCTGAAAAGAGAACTCCCCAGG 408
 |||||
 878 AA-CCCATGAACAGTATGAAAAGGCAAAATGATAGGATCTGAAAAGGAGAACTCCNAGG 820
 |||||
 409 TCAGTAGTGCCCATATGCTACTGGAGATCAGTGGAGAATACTCCAGAAGAAATGAA 468
 |||||
 819 TCAGTAGTGCCCATATGCTACTGGAGATCAGTGGAGAATACTCCAGAAGAAATGAA 760
 |||||
 469 GAGATGGAGCCAAAGCAAAAAGATATCCAGCTCTGGATGTGACTGTGTATATAAGCAAG 528
 |||||
 759 GAGATGCAGCTAAAGCAAAAACATATCCAGCTCTGGATGTGACTGTGTGTAGAGCAAG 700
 |||||
 529 GTCCGATGCTTAAGAGCAATATTCATAGAACCTTGGATGTGACTGTGTGTATCAATCAA 588
 |||||
 699 GTCCGATGCTTAAGAGCAATATTCATAGAACCTTGGATGTGACTGTGTGTATCAATCAA 640
 |||||
 589 GGCAAAATGGAAAGTGGTCAAAACAGAGATGCCAAGAGTGAATGTCAACATTTAGGAATC 648
 |||||
 639 GGCAAAATGGAAAGTGGTCAAAACAGAGATGCCAAGAGTGAATGTCAACATTTAGGAATC 580
 |||||
 649 AGCGAATAAATGAGTGGAAATGGTGAATTTAACTCAGATGACCATATATCTACTAC 708
 |||||
 579 AGCAAACTAAATGAGTGGAAATGGTGAATTTATCTCAGATGACCATTTGTATCTACTAC 520
 |||||
 709 TCGGGCAGGATCCCTCAGAGAAATGGAGTGGATCCATCATGTCAACAAAGAGTCCGA 768
 |||||
 519 TCGGGCAGGATCCCTTAGAAGAACTGGATGGATCCATCATGTCAACAAAGAAATCAA 460
 |||||
 769 AATGAGTACTTGGATGAGTCTCAAAACGACAGAAATGATCTCTGTTGTTTCCAAAGGC 828
 |||||
 459 AATGAGTACTTGGATGAGTCTCAAAACGACAGAAATGATCTCTCTGTTTCCAAAGGC 400
 |||||
 829 AAACCAATTAATACAGTAAATCAAGTCTATGCCCCCAACAGTAATGCTGAAGAGCT 888
 |||||
 399 AAACCAATTAATACAGTAAATCAAGTCTATGCCCCCAACAGTAATGCTGAAGAGCT 340
 |||||
 889 GAAGTTGAACGGTCTCTATGAAGACCTCAAGACCTTTTAGAACTAAACACCCCAAAAAGAT 948
 |||||
 339 GAAGTTGAACGGTCTCTATGAAGACCTCAAGACCTTTTAGAACTAAACACCCCAAAAAGAT 280
 |||||
 949 GTCTTCTCAATATAGGGGACTGGAATGCAAAAGTAGGAAGCAAGAAACACCTGGAGTA 1008
 |||||
 279 GTCTTCTTCAATATAGGGGACTGGAATGCAAAAGTAGGAAGTCA---AACACATGGAGTA 223
 |||||
 1009 ACAGGCAAAATTTGGCTTTGGATACGGAATGAAGCAGGGCAAGAACTAATAGAGTTTTCG 1068
 |||||
 222 ACAGGCAAAATTTGGCTTTGGATACGGAATGAAGCAGGGCAAGAACTAATAGAGTTTTCG 163
 |||||

QY 1069 CAAGAAATGCACTGGTTCATAGCAAAACCCCTCTTCCAAACAACAAGAGAAGCTTAC 1128
 |||||
 DB 162 CAAGAAATGCACTGGTTCATAGCAAAACCCCTCTTCCAAACAACAAGAGAAGCTTAC 103
 |||||
 QY 1129 ACATGACATCACCAGATGGTCAACACCGAATCAGATTGATTATTTTCGAGCC-A 1187
 |||||
 DB 102 ACATGACATCACCAGATAGTTAACTGAAATCAGACTGATTATTTTCGAGCCNA 43
 |||||
 QY 1188 AAGATGGAGAAGCTCTATACAGTCAAGCAAAAAACAAGACCAG 1228
 |||||
 DB 42 AAGATGGAGAAGCTCTATACAGTCAACAAAAACAAGACCAG 2

RESULT 13

BZ882580

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BZ882580 922 bp DNA linear GSS 18-MAR-2003
 CH240_195M2_TV3 CHORI-240 Bos taurus genomic clone CH240_195M2,
 genomic survey sequence.

BZ882580
 BZ882580.1 GI:29109982
 GSS.

Bos taurus (cow)
 Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.

1 (bases 1 to 922)
 Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,
 Shvartsbeyn,A., Gebregorgis,E., Chen,D., Riggs,F., de Jong,P.,
 Crawford,A.M. and McEwan,J.C.
 Bovine BAC End Sequences from Library CHORI-240
 Unpublished (2003)

Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200

Fax: 301 838 0208
 Email: szhao@tigr.org

Clones are derived from the bovine BAC library CHORI-240
 (http://www.chori.org/bacpac/bovine240.htm). For BAC library
 availability, please contact Pieter de Jong (pdejong@mail.cho.org).
 Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/ordering/information.html). This work
 was undertaken as part of the International Bovine BAC Mapping
 Consortium (IBBMC) by AgResearch Ltd., New Zealand and The
 Institute of Genomic Research (TIGR), USA.
 Plate: 195 row: M column: 2
 Seq primer: T7

Class: BAC ends.

FEATURES

Location/Qualifiers
 1. .922
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /strain="breed: Hereford"
 /db_xref="taxon:9913"
 /clone="CH240_195M2"
 /sex="Male"
 /cell_type="Blood"
 /clone_lib="CHORI-240"
 /note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
 Hereford bull Li Domino 99375; CHORI-240 Bovine BAC
 library (Male) produced by Pieter de Jong"

Query Match 22.3%; Score 797.8; DB 1; Length 922;
 Best Local Similarity 94.8%; Pred. No. 0.14; Mismatches 42; Indels 8; Gaps 5;
 Matches 877; Conservative 0;

QY 1692 ACATCTATTCTCTTTATGACTGCAAAAGCCCTTGGCTGGGGTCAATAAAC 1751
 |||||
 DB 1 ACACTCTTCTCTTTATGACTATGCCAAAGCCCTTGGCTGGGTGATCAATAAAC 60
 |||||

QY 1752 TGTGAAATTTCTGAAAGGATGGGAATACAGAGCCACCTGACCTGACTCTTGAAAAATT 1811
 Db |||||
 61 TGTGAAATTTCTGAAACAAATGGGAATACCAGAGCCACCTGACCTGCTTGTGAGAAC 120
 QY 1812 TGTATGTCAGGTGAGGAAGCAACAGTGTAGAACTGGACATGGCAACACAGACTGGTTTCAAG 1871
 Db |||||
 121 TATATGTCAGGTGAGGAAGCAACAGTGTAGAACTGGACATGGCAACACAGACTGGTTTCAAA 180
 QY 1872 TAGGAAAGAGGTATGTCAAGGCTATATTTGTCAACCGGCTGTTTAACTTCTATGAG 1931
 Db |||||
 181 TAGGAAAGAGGTATCATCAAGGCTATATTTGTCAACCGCTGTTTAACTTCTATGAG 240
 QY 1932 AG-ACATCATGAGAAACCTGGCTGGGAAGAGACACAGCTGGGAATCAAGATTCGCGGA 1990
 Db |||||
 241 AGTACATCATGAGAAATGTTGGCTGGGAAGAGACACAGCTGGGAATCAAGACTGCTGGGA 300
 QY 1991 GAAATAGCAATAACCTCAGATATGAGATGATACACACCTTTATGGCAGAAAGTGAAGAG 2050
 Db |||||
 301 GAAATATCAATAACCTCAGATACGAGATGATACACACCTTTATGGCAGAAAGTGAAGAG 360
 QY 2051 AACTTAAAGAGCTCTTGATGAGGTGAAAGGAGAGTGAAGAGTGGCTTAAAGTCA 2110
 Db |||||
 361 AACTTAAAGAGCTCTTGATGAGGTGAAAGGAGAGTGAAGAGTGGCTTAAAGTCA 420
 QY 2111 ACATTCAGAAACGAGATCATGTCATCTGCTGCCATCACCTTCATGGGAATAGATGGG 2170
 Db |||||
 421 ACATTCAGAAACGAGATCATGTCATCTGCTGCCATCACCTTCATGGGAATAGATGGG 480
 QY 2171 AAACAGTGAACAGTGTGAGCTTATTTTGTGGGGCTTCCAAAATCACTGCAGATGG 2230
 Db |||||
 481 AAACAGTGAACAGTGTGAGCTTATTTTGTGGGGCTTCCAAAATCACTGCAGATGG 538
 QY 2231 TGACTGCGAGCCATGAAATTTAAAGACACTTACTCTCTTGGAGAGAAAGTTA--ACCAACCT 2288
 Db |||||
 539 TGACTGCGAGCCATGAAATTTAGAGACGCTTACTCTTGGAGAGAAAGTTATGACTTAACCT 598
 QY 2289 AGATAGCATATTGAAAGCAGAGACATTAACCTTCCCAACAAAGCCCATCTAGTCAAGGC 2348
 Db |||||
 599 AGATAGCATATTGAAAGCAGAGACATTAACCTTCCCAACAAAGCTCCATCTAGTCAAGGC 658
 QY 2349 TATGTTTTTCCAGTGGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2408
 Db |||||
 659 TATGTTTTTCCAGTGGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 718
 QY 2409 CTGAGAAATTGATGTTTTGAACTGTGTTGAGAGAGACTCTTGTAGAGTCCCTTGGAC 2468
 Db |||||
 719 CCGAAGAAATTGATGTTTTGAACTGTGTTGAGAGAGACTCTTGTAGAGTCCCTTGGAC 778
 QY 2469 TCGAAGAGATCCACAGTCCATCTGAGAGATCAGCCCTGGGATTTCTTGGAGG 2528
 Db |||||
 779 TCGAAGAGATCCACAGTCCATCTGAGAGATCAGCCCTGGGATTTCTTGGAGG 838
 QY 2529 AATGATGCTAAAGCTGAAACTCCAGTACTTTGGCCACCTGATCAGAGAGCTGACTCACT 2588
 Db |||||
 839 AATGATGCTAAAGCTGAAACT--CAGTACTTTGGCCACCTCATGCGAGAGTTGACTCACT 897
 QY 2589 GGAAAGACCCCTGATGCTGGAGGGAT 2615
 Db |||||
 898 GGAAAGA--TTGATGCTGGAGGGAT 922

RESULT 14
 CC480115/c
 LOCUS CH240_30704.77 CHORI-240 Bos taurus genomic clone CH240_30704,
 DEFINITION genomic survey sequence.
 ACCESSION CC480115
 VERSION CC480115.1 GI:31759328
 KEYWORDS GSS.
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

REFERENCE AUTHORS

Bovidae; Bovinae; Bos.
 1 (bases 1 to 899)
 Holt,R., Scott,J., Yang,G., Barber,S., Smailus,D., Prabhu,A.-L.,
 Tsai,M., Cloutier,A., Lee,D., Girm,N., Olson,T., Mayo,M.,
 Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R.,
 Mathewson,C., Wye,N., Masson,A., Brown-John,M., Jones,S.,
 Schein,J., Marra,M., de Jong,P., McWilliam,S., Barris,W.,
 Dalrymple,B.P. and Tellam,R.
 Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
 Unpublished (2003)
 Other GSSs: CH240_30704.TARBAC13P2
 Contact: Rob Holt
 Sequencing
 The British Columbia Cancer Agency Genome Science Centre
 600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
 Tel: 604-877-6085
 Fax: 604-877-6276
 Email: rholt@bcgsc.ca
 Clones are derived from the bovine BAC library CHORI-240
 (http://www.chori.org/bacpac/bovine240.htm). For BAC library
 availability, please contact Pieter de Jong (pdejong@mail.cho.org).
 Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/ordering/information.htm). This work
 was undertaken as part of the International Bovine BAC Mapping
 Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
 British Columbia Genome Sciences Centre, Canada.
 Plate: 307 row: O column: 4
 Seq primer: T7
 Class: BAC ends.
 Location/Qualifiers
 1..899
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /strain="breed: Hereford"
 /db_xref="taxon:9913"
 /clone="CH240_30704"
 /sex="Male"
 /cell_type="Blood"
 /clone_lib="CHORI-240"
 /note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
 Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
 library (Male) produced by Pieter de Jong"

TITLE JOURNAL COMMENT

FEATURES source

Query Match 22.3%; Score 797.2; DB 1; Length 899;
 Best Local Similarity 94.5%; Pred. No. 0.15;
 Matches 826; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
 725 TCAGAGAAATGAGTAGCCATCATGGTCAACAAAGAGTCCGAATGCAGTCTGGAT 784
 Db |||||
 899 TCAGAGAAATGAGTAGCCATCATGGTCAACAAAGAGTCCGAATGCAGTCTGGAT 840
 QY 785 GCAGTCTCAAAACGACAGATGATCTCTGTTGTTTCCAAAGGCAACCATCAATATCA 844
 Db |||||
 839 GCAATCTCAAAATGACAGACAATATCTGTTTCGTTTCCAAAGGCAACCATCAATATCA 780
 QY 845 CAGTATCAAGTCTATGCCCAACAGTAACTGCTGAGAGAGTCAAGTGAACGCTCT 904
 Db |||||
 779 CAGTATCAAGTCTATGCCCAACAGTAACTGCTGAGAGAGTCAAGTGAACGCTCT 720
 QY 905 ATGAGACCTCAAGACCTTTTAGAACTAAACCCCAAAAAGATGCTCTTCTCATTATAG 964
 Db |||||
 719 GTGAGAACTCAAGACCTTTTAGAACTAAACCCCAAAAAGTGTCTCTTCTATTATAG 660
 QY 965 GGGACTGGAAATGCAAAAGTAGGAAGCAAGAAACACCTGGAGTAAACGGCAAAATTGGCC 1024
 Db |||||
 659 GGGACTGGAAATGCAAAAGTAGGAAGTCAAGAAACACCTGGAGTAAACGGCAAAATTGGCC 600
 QY 1025 TTGGAATACGGAATGAAGCAGGGCAAGAGTAACTAATAGAGTTTGCACAGAAATGCACTGG 1084
 Db |||||
 599 TTGGAATACGGAATGAAGCAGGGCAAGAGTAACTAATAGAGTTTGTCTAAGAAATGCACTGG 540
 QY 1085 TCATAGCAAAACCTCTTCCAAACCAACAGAGAGAGTCTTACACATGGACATCACCAG 1144
 Db |||||

Db 539 TCATAAACAACCCCTCTTCCACAACAACAAGAGAAGACTATATATCATGAGCATCAAA 480

QY 1145 ATGTCACACCGAATTCAGATTGATTATATCTTTTCAGCCCAAGAGTGAAGAGCTCTA 1204

Db 479 ATGTCACACCGAATTCAGATTGATTATATCTTTTCAGCCCAAGAGTGAAGAGCTCTA 420

QY 1205 TAGAGTCAGCAAAACAAGACCGAGGAGCTTACTGTGGCTCAGATCATGAATCTTTATTG 1264

Db 419 TAGAGTCAGCAAAACAAGACCGAGGAGCTGACCGTGGTTTCAGACCATGAATCTTTATTG 360

QY 1265 CCAAAATTCAGACTTAAATTCAGAAAGTAGGAAACCACTAGATCACTCAGTGAAGACC 1324

Db 359 CCAAAATTCAGACTTAAATTCAGAAAGTAGGAAACCACTAGATCACTCAGTGAAGACC 300

QY 1325 TAAATCCAAATCCCTTATGATTATACAGTGGAAAGTGAATAAGATTAAGGGCTAGATC 1384

Db 299 TAAATCCAAATCCCTTATGATTATACAGTGGAAAGTGAATAAGATTAAGGGCTAGATC 240

QY 1385 TGATAGCAGAGTACCTTAATGAACTATGACAGAGGTTTCATGACATGTGTACAGGAGACAG 1444

Db 239 TGATAGCAGAGTACCTTGAATGAACTATGAAATGAGGTTTCATGACATGTGTACAGGAGACAG 180

QY 1445 GGATCGAGACATCCCATCGAAAGAAATGCAAAAAGCAAAATGGCTCTCTGGGAGG 1504

Db 179 GGATCGAGACATCCCATCGAAAGAAATGCAAAAAGCAAAATGGCTCTCTGGGAGG 120

QY 1505 CTTTACAAATAGCTGTGAAAGAGAGAGTGAAGCAAGCAAGAAAGAAAGAAAGTAA 1564

Db 119 CTTTACAAATAGCTGTGAAAGAGAGAGTGAAGCAAGCAAGAAAGAAAGAAAGTAA 60

QY 1565 AGCATCTGAATGCAGAGTTCCAAGAACTTCCAG 1598

Db 59 AACATCTGAATGCAGAGTTCCAAGAAATAGCAAG 26

RESULT 15

CC475230

LOCUS

DEFINITION CC475230 958 bp DNA linear GSS 16-JUN-2003
genomic survey sequence.

ACCESSION

CC475230

VERSION

CC475230.1

KEYWORDS

GSS.

SOURCE

Bos taurus (cow)

ORGANISM

Bos taurus

REFERENCE

AUTHORS

Holt, R., Stott, J., Yang, G., Barber, S., Small, D., Prabhu, A.-L.,
Tsai, M., Cloutier, A., Lee, D., Girn, N., Olson, T., Mayo, M.,
Butterfield, V., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
Mathewson, C., Wye, N., Masson, A., Brown, J., Jones, S.,
Schein, J., Marra, M., de Jong, P., McWilliam, S., Barris, W.,
Dairymple, B. P. and Tellam, R.

TITLE

Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398

JOURNAL

Unpublished (2003)

COMMENT

Other GSSs: CH240_300L11.TARBAC13P2
Contact: Rob Holt
Sequencing
The British Columbia Cancer Agency Genome Science Centre
600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4S6
Tel: 604-877-6085
Fax: 604-877-6276
Email: rholt@bcgsc.ca

Clones are derived from the bovine BAC library CHORI-240
(<http://www.chori.org/bacpac/bovine240.htm>). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(<http://www.chori.org/bacpac/ordering/information.htm>). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
British Columbia Genome Science Centre, Canada.

Plate: 300 row: L column: 11

Seq primer: T7

Class: BAC ends.

FEATURES

source

Location/Qualifiers
1..958
/organism="Bos taurus"
/mol_type="genomic DNA"
/strains="bred: Hereford"
/db_xref="taxon:9913"
/clone="CH240_300L11"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
Hereford bull LI Domino 99375; CHORI-240 Bovine BAC
library (Male) produced by Pieter de Jong"

Query Match 22.3%; Score 796.9; DB 1; Length 958;

Best Local Similarity 92.8%; Pred. No. 0.14;

Matches 887; Conservative 0; Mismatches 62; Indels 7; Gaps 5;

QY 1748 AAACCTGTGGAAATCTCAAGGGATGGGAATACAGACCACCTGACCTGACTCTTGAA 1807

Db 6 AACTGTGTGGAAATCTCTCAAGAGATGGGAATACAGACCACCTGACCTGCTCTTGAGA 65

QY 1808 AATTTGTATGCAGGTCAGGAACCAACAGTTAGAACTGGACATGGAAACAACAGACT -GGTT 1866

Db 66 AATCTGTATGCAGGTCAGGAACCAACAGTTAGAACTGGACATGGAAACAACAGACTCGGTT 125

QY 1867 CCAAGTAGGAAAGAGTAGTATGCAAGGCTGTATATGTCACCCGGCTTGTTAATCTCTA 1926

Db 126 CCAAAACAGGAAAGAGTAGCTCAAGGCTGTATATGTCATCTCTTATTTAATTTATA 185

QY 1927 TGCAGAG -ACATCATGAGAAACGCTGGGCTGGAAGACACAAAGCTGGAATCAAGATTGC 1985

Db 186 TGCAGAGTACATCATGAGAAATGCTGGCTGGAAGACACAAAGCTGGAATCAAGATTGC 245

QY 1986 CCGGGAATAATCAACAATCACTCAGATATGCAGATGATACCAACCTTATGGCAGAAAGTGA 2045

Db 246 CAGGGGAATATCAATAATCTCAGATATGCAGATGATACCAACCTTATGGCAGACAGTGA 305

QY 2046 AGAGGAACTAAAAGGCTCTTGCATGAAGTGAAGAGGAGTGAAGAAAGTTGGCTTAAA 2105

Db 306 AGAGGAACTAAAAGGCTCTTGGTGAAGTGAAGAGTGAAGAAAGTTGGCTTAAA 365

QY 2106 GGTCAACATTCAGAAAACGAAGATCATGCATCTGGTCCCATCACTTCATGGGAAATAGA 2165

Db 366 GGTCAACATTCAGAAAATGAAGATCATGCATCTGGTCCCATCACTTCATGGGAAAGAGA 425

QY 2166 TGGGGAACAGTGGGAAACAGTGTGAGCTTTATTTTGGGGGGCTCCAAAATCACTGCA 2225

Db 426 TGGGGAACAGTGGGAAACAGTGTGAGCTTTATTTTCTGGGCTCCAAAATCACTGCA 483

QY 2226 GATGGTGACTCGAGCCCATGAAATTAAGACACACTTACTCTTTGGAAGAAAAGTTA -ACC 2283

Db 484 GATGGTGACTCGAGCCCATGAAATTAAGAGGCGCTTACTCTTTGGAAGAAATGTTATGACC 543

QY 2284 AACCTAGATAGCATATTGAAAACGAGAGACATTAACCTGCCAACAAAGCCCATCTAGTC 2343

Db 544 AACCTAGATAGCATATTGAAAACGAGAGACATTAACCTGCCAACAAAGGTCCTCTAGTC 603

QY 2344 AAGGCTATGGTTTTTCCAGTGGTCATGTATGATGTGAGAGTTGGAGCTGTGAAGAAAGCT 2403

Db 604 AAGTCTATGGTTTTTCCAGTGGTCATGTATGATGTGAGAGTTGGAGCTGTGAAGAAAGCT 663

QY 2404 GAGCACTGAAGAAATTCATGCTTTTGAACCTGTGGTGTGGAGAGAGACTCTTGAGAGTCCCT 2463

Db 664 GAGCACTGAAGAAATTCATGCTTTTGAACCTGTGGTGTGGAGAGAGACTCTTGAGAGCCCT 723

QY 2464 TGAAGTGAAGAGATCCAAACAGTCCATCTGAAGAGATCAGCCCTGGATTTCTTTG 2523

Db 724 TGAAGTGAAGAGAT -CAACCAAGTCCATTTAAAGAGATCAGTCTCTGGGTGTTCTTTG 782

availability, please contact Pieter de Jong (pdejong@mail.cho.org).

Clones may be purchased from BACPAC Resources
(http://www.choi.org/bacpac/ordering_information.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
British Columbia Genome Sciences Centre, Canada.

Plate: 415 row: B column: 11

Seq primer: SP6

Class: BAC ends.

FEATURES

Location/Qualifiers

1..904
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="bred: Hereford"
/db_xref="taxon:9913"
/clone="CH240_415B11"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
library (Male) produced by Pieter de Jong"

Query Match 22.1%; Score 790.8; DB 1; Length 904;
Best Local Similarity 94.7%; Pred. No. 0.15;
Matches 859; Conservative 0; Mismatches 42; Indels 6; Gaps 4;

QY	289	TCAGAGATGTGTCCTAGGAGAA--GGGATCGCAACACCTTCAGTATCTTGGCTTG	346
DB	1	TCAGAGATGTGTCCTAGGAGAA--GGGATCGCAACACCTTCAGTATCTTGGCTTG	60
QY	347	AGAACCCCATCAACAGTATGAAAGGCAAAATAGTAGTACTGAAAGAGAACTCCCCA	406
DB	61	AGAACCCCATCAACAGTATGAAAGGCAAAATAGTAGTACTGAAAGAGAACTCCCTCA	120
QY	407	GTCAGTAGTGGCCCATGCTACTGGAGATCAGTGGAGAAATATCTCAGAAAGATG	466
DB	121	GATCAGTAAGGGCCCAATAGTCTACTGGAGATCGGTGGAGAAATATCTCAGAAAGATG	180
QY	467	AAGAGATGGAGCCCAAGCAAAAGAAATATCCAGCTGTGGATGTGACTGGTGATATAGCA	526
DB	181	AGGGATGGAGCCCAAGCAAAAGAAATATCCAGCTGTGGATGTGACTGGTGATATAGCA	240
QY	527	AGTCCGATGCTGTAAGAGCAATATTGTCATAGGAACCTCGAATGTGAGTCCATGAATC	586
DB	241	AGTCCGATGCTGTAAGAGCAATATTGTCATAGGAACCTCGAATGTGAGTCCATGAATC	300
QY	587	AAGCAAAATGGAGTGTCAACACAGATGCGAGAGTGAATGTCAACATTTCTAGGA	646
DB	301	AAGCAAAATGGAGTGTCAACACAGATGCGAGAGTGAATGTCAACATTTCTAGGA	360
QY	647	TCAGCAACTAAATGGACTGGAATGGTGAATTTAACTCAGATGACCATTAATCTACT	706
DB	361	TCAGCAACTAAATGGACTGGAATGGTGAATTTAACTCAGATGACCATTAATCTACT	420
QY	707	ACTCGGGCAGGATCCCTCAGAGAAATGAGTAGGCATCATGTCACAAAGAGTCC	766
DB	421	ACTCGGGCAGGATCCCTCAGAGAAATGAGTAGGCATCATGTCACAAAGAGTCC	478
QY	767	GAATGCAGTACTGGATGAGTCTCAAAAACGACAGATGATCTCTGTTGTTTCCCAAG	826
DB	479	GAATGCAGTACTGGATGAGTCTCAAAAACGACAGATGATCTCTCTCTGTTTCCCAAG	538
QY	827	GCAACCAATTCATATCAGATGATGATGATGATGATGATGATGATGATGATGATGATG	886
DB	539	GCAACCAATTCATATCAGATGATGATGATGATGATGATGATGATGATGATGATGATG	598
QY	887	CTGAAGTTGAACCGTCTCTATGAAGACTCAACAGCTTTTAGAAGTAAACACCCCAAAAG	946
DB	599	CTGAAGTTGAACCGTCTCTATGAAGACTCAACAGCTTTTAGAAGTAAACACCCCAAAAG	658
QY	947	ATGTCCTTCTCATATAGGGAGTGGAAATGCAAAAGTAGGAAGCAAAAGAACCTTGGAG	1006

DB	659	ATGTCCTTTTCATTATATAGGGAGTGGAGTGCAAAAGTAGGAAGTCAAGAAACACCTGGAG	718
QY	1007	TAACAGGCAAAATTTGGCTTTGGAATACGAATGAAGCAGGCAAGAGACTAAT-AGAGTTT	1065
DB	719	TAACAGGCAAAATTTGGCTTTGGAATACGAATGAAGCAGGCAAGAGACTAAT-AGAGTTT	778
QY	1066	TGCCAAGAAAATGCACTGTGTATAGCAAAACACCTCTTCCCAACACCAAGAGAGACTC	1125
DB	779	TGCCAAGAAAATGCACTGTGTATAGCAAAACACCTCTTCCCAACACCAAGAGAGACTC	838
QY	1126	TACACATGACATCACCAGATGTCACACCCCAATACAGATGATATATTTTTCGAGC	1185
DB	839	TATACATGACATCACCAGATGTCACACCCCAATACAGATGATATATTTTTCGAGC	897
QY	1186	CAAAGAT 1192	
DB	898	CAAAGAT 904	

CC476492 897 bp DNA linear GSS 16-JUN-2003
CH240_30208.T7 CHORI-240 Bos taurus genomic clone CH240_30208,
genomic survey sequence.

ACCESSION CC476492
VERSION CC476492.1 GI:31753755
KEYWORDS GSS.

SOURCE Bos taurus (cow)

ORGANISM

REFERENCE 1 (bases 1 to 897)
Holt, R., Stott, J., Yang, G., Barber, S., Smailus, D., Prabhu, A., L.,
Tsai, M., Cloutier, A., Lee, D., Girm, N., Olson, T., Mayo, M.,
Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
Schein, J., Marra, M., de Jong, P., McWilliam, S., Barris, W.,
Dalrymple, B. P. and Tellam, R.

AUTHORS

TITLE

JOURNAL

COMMENT

Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
Unpublished (2003)
Other_GSSs: CH240_30208.TARBAC13P2
Contact: Rob Holt

Sequencing

The British Columbia Cancer Agency Genome Science Centre

600 W. 10th Ave. Vancouver, British Columbia, Canada V5Z 4E6

Tel: 604-877-6085

Fax: 604-877-6276

Email: rholt@bccsc.ca

Clones are derived from the bovine BAC library CHORI-240

(<http://www.choi.org/bacpac/bovine240.htm>). For BAC library

availability, please contact Pieter de Jong (pdejong@mail.cho.org).

Clones may be purchased from BACPAC Resources

(http://www.choi.org/bacpac/ordering_information.htm). This work

was undertaken as part of the International Bovine BAC Mapping

Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the

British Columbia Genome Sciences Centre, Canada.

Plate: 302 row: O column: 8

Seq primer: T7

Class: BAC ends.

FEATURES

Location/Qualifiers

1..897
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="bred: Hereford"
/db_xref="taxon:9913"
/clone="CH240_30208"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
library (Male) produced by Pieter de Jong"

Query Match		22.1%;	Score 790.2;	DB 1;	Length 897;
Best Local Similarity		95.1%;	Pred. No. 0.16;		
Matches 856;		Conservative	0;	Mismatches	38; Indels 6; Gaps 4;
QY	1664	ATGGAAAAACAAGAGAGTTCAGAAAAACATCTATTTCCTTTATTGACTATGCAAAA	1723		
DB	1	ATGGAAAAACAAGAGAGTTCAGAAAAACCATCTATTTCCTGCTTTATTGACTATGCCCCAAA	60		
QY	1724	GCCTTTGACTGTGGGGGTCCACAATAAACTGTGAAAAATTTCTGAAAGGGATGGGAATACCA	1783		
DB	61	GCCTTTGACTGTGTGGATCACAACTGTGAAAAATTTCTGAAAGAGATAGGAATACCA	120		
QY	1784	GACCACCTGACCTGACTCTTTGAAAAATTTGTATGACAGTCAGGAGCAACAGTTAGAACT	1843		
DB	121	GACCACCTGATCTGCTCTTTGAGAAATCTGTATGACAGTTCAGGAGCAACAGTTAGAACT	180		
QY	1844	GGACATGGAAACAACAGACTGGTTCCAAGTAGGAAAAAGGAGTAGTGTCAAGGGCTGTATATTG	1903		
DB	181	GGACATGGAAACAACAGACTGGTTCCAATATAGGAAAAAGGACTTCGTCAAGGCTGTATATTG	240		
QY	1904	TCACCCGGCTGTTTAACTTCTATGCAGAG-ACATCATGAGAAAACTGGGCTGGAAGAA	1962		
DB	241	TCACCCGTGCTTATTTAAATTTATATGCAGAGTAGCATCATGAGAAAACTGGGCTGGAAGAA	300		
QY	1963	GCACAGCTGGAATCAAGATTTCGCGGAGAAATAGCAATAACCTCAGATATGCAGATCAT	2022		
DB	301	GCACAGCTTGGAAATCAAGATTTCGCGGAGAAATATCAATAACCTCAGATATGCAGATCAT	360		
QY	2023	ACCACCTTTATGGCAGAAAGTGAAGAGCAATTAAGAGCCTTTGATGAAGTGAAGAG	2082		
DB	361	ACCACCTTTATGGCAGAAAGTGAAGAGCAATTAAGAGCCTTTGATGAAGTGAAGAG	420		
QY	2083	GAGAGTGAAGAAAGTGGCTTAAAGCTCAACATTCAGAAAAACGAGATCATGGCATCTGCT	2142		
DB	421	GAGAGTGAAGAAAGTGGCTTAAAGCTCAACATTCAGAAAAACGAGATCATGGCATCTGCT	480		
QY	2143	CCCATCTCTCATGGGAAATAGATGGGGAACAGTGGAAACAGTGTCCAGACTTTATTTTT	2202		
DB	481	CCCATCTCTCATGGGAAATAGATGGGGAACAGTGGAAACAGTGTCCAGACTTTATTTTT	540		
QY	2203	TGGGGGGCTCCAAAATCACTGCAGATGGTCACTGCAGCCCATGAAATTAAGACACACTTAC	2262		
DB	541	--CTGGGCTCCAAAACCACTGCAGATGGTCACTGTAGCCATGAAATTAAGACACACTTAC	598		
QY	2263	TCCTTGGAGAAAAAGTTA--ACCAACTAGATAGCATATTGAAAAACAGACACATTACCT	2320		
DB	599	TCCTTGGAGAAAAAGTTATGACCAACTAGATAGCATATTCAAAAGCAGACACATTACTT	658		
QY	2321	TGCCAACAAAGCCCCATCTAGTCAAGGCTATGGTTTTTCCAGTGGTCACTATGAGATGTG	2380		
DB	659	TGCCAACAAAGGTTGGTCTAGTCAAGGCTATGGTTTTTCCAGTGGTCACTATGAGATGTG	718		
QY	2381	AGAGTTGGACTGTGAAGAAAGCTGAGCACTGGAAGAAATTGATGCTTTTGAAGTGGTGTT	2440		
DB	719	AGAGTTGGACTGTGAAGAAAGCTGAGCACTGGAAGAAATTGATGCTTTTGAAGTGGTGCT	778		
QY	2441	GGAGAGACTCTTTGAGAGTCCCTTGGACTGCCAAGGAGATCCAAACAGTCCATTCTGAAGG	2500		
DB	779	GGAGAGACTCTTTGAGAAATCCCTTGGACTGCCAAGGAGATCCAAACAGTCCATTCTGAAGG	838		
QY	2501	AGATCAGCCCTGGGATTTCTTTTGGAGGAATGATGCTTAAAGCTGAAACTCCAGTACTTTG	2560		
DB	839	AGATCAGTCTCTGGGATTTCTTTTGGAA--GAATGATGCTTAAAGCTGAAACTCCAGTACTTTG	897		
RESULT 20					
CC579657/c					
LOCUS					
DEFINITION					
CH240_374G1.T7 CHORI-240 Bos taurus genomic clone CH240_374G1,					
genomic survey sequence.					
ACCESSION					
CC579657					
VERSION					
CC579657.1 GI:31925542					

KEYWORDS	GSS.
SOURCE	Bos taurus (cow)
ORGANISM	Bos taurus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
AUTHORS	1 (bases 1 to 881) Holt, R., Stott, J., Yang, G., Barber, S., Smailus, D., Prabhu, A.-L., Tsai, M., Cloutier, A., Lee, D., Girn, N., Olson, T., Mayo, M., Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R., Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S., Schein, J., Marra, M., de Jong, P., McWilliam, S., Barrie, W., Dalrymple, B. P. and Telling, R.
TITLE	Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
JOURNAL	Unpublished (2003)
COMMENT	Other_GSSs: CH240_374G1.TARBAC13P2 Contact: Rob Holt Sequencing The British Columbia Cancer Agency Genome Science Centre 600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4B6 Tel: 604-877-6085 Fax: 604-877-6276 Email: rholt@bcgsc.ca Clones are derived from the bovine BAC library CHORI-240 (http://www.chori.org/bacpac/bovine240.htm). For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/ordering information.htm). This work was undertaken as part of the International Bovine BAC Mapping Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the British Columbia Genome Sciences Centre, Canada. Plate: 374 row: G column: 1 Seq primer: T7 Class: BAC ends.
FEATURES	Location/Qualifiers 1.. 881 /organism="Bos taurus" /mol_type="genomic DNA" /strain="breed: Hereford" /db_xref="taxon:9913" /clone="CH240_374G1" /sex="Male" /cell_type="Blood" /clone_lib="CHORI-240" /note="Vector: PTARBAC1.3; Site 1: MboI; Site 2: MboI; Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC library (Male) produced by Pieter de Jong"
Query Match	22.0%; Score 785.2; DB 1; Length 881;
Best Local Similarity	95.0%; Pred. No. 0.17;
Matches 811; Conservative	0; Mismatches 43; Indels 0; Gaps 0;
Qy	745 ATCATGGTCAACAAAAGAGTCCGAAATGCAGTACTTGGATGCAGTCTCAAAAACGACAGA 804
Db	881 ATCATGGTCAACAAAAGAGTCCGAAATGCAGTACTTGGATGCAGTCTCAAAAACGACAGA 822
Qy	805 ATGATCTCTGTTGTTTCCAAAGCAACCAATTCATATCACAGTATCCAAAGTCTATGCC 864
Db	821 ATGATCTCTGTTGTTTCCAAAGCAACCAATTCATATCACAGTATCCAAAGTCTATGCC 762
Qy	865 CCAACCAAGTATGCTCAAGAGCTGAAGTTGAACGGTCTTATGAAGACCTTACAGACCTT 924
Db	761 CCAACCAAGTATGCTCAAGAGCTGAAGTTGAACGGTCTTATGAAGACCTTACAGACCTT 702
Qy	925 TTAGAACTAACCCCAAAAAGATGTCCTTCTCATTTATAGGGGACTGGAATGCAAAAGTA 984
Db	701 TTAGAACTAACCCCAAAAAGATGTCCTTCTCATTTATAGGGGACTGGAATGCAAAAGTA 642
Qy	985 GGAAGCAAGAAACACCTGGAGTAAACAGGCAAAATTTGGCTTTGGAAATACGGAAATGAAGCA 1044
Db	641 GGAAGTCAAGAAACACCTGGAGTAAACAGGCAAAATTTGGCTTTGGAAATACAGAAATTAAGCA 582
Qy	1045 GGGCAAAAGACTAATAGAGTTTTGCCAAGAAATGCAGTCTGTCATAGCAAAACACCCCTCTTC 1104

```

Db      581 GGGCAAGACTAATAGAGTTTGCCAGAAATGACCTGATCATAAACAAACACCTCTTC 522
QY      1105 CAACAACAAGAGAGAGACTCTACATGGACATCACAGATGTCACACCGAAATCAG 1164
Db      521 CAACAACAAGAGAGAGACTCTATACATGGATATCAACAGATGTCACACCGAAATCAG 462
QY      1165 ATTGATTATATCTTTGAGCCAAAGATGGAGAGCTCTATACAGTACGAAAAACAGA 1224
Db      461 ATTGATTATATCTTTGAGCCAAAGATGGAGAGCTCTATACAGTACGAAAAACAGA 402
QY      1225 CCAGGAGCTTACTGTGGCTCAGATCATGAATCTCTATTGCAAAATTCAGACTTTAAATTG 1284
Db      401 CCAGGAGCTTACTGTGGCTCAGACCATGAATCTCTATTGCAAAATTCAGACTTTAAATTG 342
QY      1285 AAGAAAGTAGGAAACCACTAGATCATCTAGGTAAGACCTTAATCCAAATCCCTTTATGAT 1344
Db      341 AAGAAAGTCGGAAACCACTAGACCATGAATCTCTATTGCAAAATTCAGACTTTAAATTG 282
QY      1345 TATACAGTGAAGTGAAGAAATAGATTTTAAGGGCTTAGATCTGTATGATACAGAGTACCTAAT 1404
Db      281 TATACAGTGAAGTGAAGAAATAGATTTTAAGGGCTTAGATCTGTATGATACAGAGTACCTAAT 222
QY      1405 GAATCTAGGACAGAGTTTCATGACATTTGATACAGGACAGGATCGAGACCATCCCATG 1464
Db      221 GAATCTAGGATGAGTTCTTTCATGACATTTGATACAGGACAGGATCGAGACCATCCCATG 162
QY      1465 GAAAGAAATCAAAAAGCAAAATGCTGTCTGGGAGGCTTACAAATAGCTGTGAAA 1524
Db      161 GAAAGAAATCAAAAAGCAAAATGCTGTCTGGGAGGCTTACAAATAGCTGTGAAA 102
QY      1525 AGAAGAGAGTCAAAAGCAAAAGCAAAATGCTGTCTGGGAGGCTTACAAATAGCTGTGAAA 1584
Db      101 AGAAGGAGTCAAAAGCAAAAGCAAAATGCTGTCTGGGAGGCTTACAAATAGCTGTGAAA 42
QY      1585 CAAGAAGTCCAG 1598
Db      41 CAAGAAGTCCAG 28

RESULT 21
CC500971
LOCUS   CH240_339C15.TARBAC13P2 CHORI-240 Bos taurus genomic clone
DEFINITION
ACCESSION CC500971
VERSION   CC500971
KEYWORDS  GSS.
SOURCE    CC500971.1 GI:31819264
          Bos taurus (cow)
          ORGANISM
          Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoides;
          Bovidae; Bovinae; Bos.
          1 (bases 1 to 925)
AUTHORS   Holt, R., Stott, J., Yang, G., Barber, S., Smalilus, D., Prabhu, A.-L.,
          Tsai, M., Cloutier, A., Lee, D., Gilm, N., Olson, R., Mayo, M.,
          Butterfield, J., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
          Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
          Schein, J., Marra, M., de Jong, P., McWilliam, S., Barris, W.,
          Dalrymple, B. P. and Tellam, R.
          Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
          Unpublished (2003)
          Other_GSSs: CH240_339C15.T7
          Contact: Rob Holt
          Sequencing
          The British Columbia Cancer Agency Genome Science Centre
          600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
          Tel: 604-877-6085
          Fax: 604-877-6276
          Email: rholt@bccsc.ca
          Clones are derived from the bovine BAC library CHORI-240
          (http://www.chori.org/bacpac/bovine240.htm). For BAC library
          availability, please contact Pieter de Jong (pdjong@mail.choi.org).

```

Clones may be purchased from BACPAC Resources (<http://www.chori.org/bacpac/ordering/information.htm>). This work was undertaken as part of the International Bovine BAC Mapping Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the British Columbia Genome Sciences Centre, Canada.

Plate: 339 row: C column: 15

Seq primer: SP6

Class: BAC ends.

FEATURES

Location/Qualifiers	1..925
/organism="Bos taurus"	
/mol_type="genomic DNA"	
/strain="breed: Hereford"	
/db_xref="taxon:9913"	
/clone="CH240_339C15"	
/sex="Male"	
/cell_type="Blood"	
/clone_lib="CHORI-240"	
/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI; Hereford bull, Li Domino 99375; CHORI-240 Bovine BAC library (Male) produced by Pieter de Jong"	

Query Match 22.0%; Score 784.6; DB 1; Length 925;
 Best Local Similarity 93.0%; Pred. No. 0.16;
 Matches 863; Conservative 0; Mismatches 59; Indels 6; Gaps 4;

QY	1796	TGACTCTTGAATAATTTGTATGACGTCAGGAGCAACAGTTAGTAACCTGGACATGGAACA	1855
Db	1	TCGCTCTTGAATAATTTGTATGACGTCATGAAGCAACAGTTAGTAACCTGGACATGGAACA	60
QY	1856	ACAGACTGTTTCCAAGTAGGAAAGAGTATGTCGAAGCTGTATATTCTCACCCTGTG	1915
Db	61	ACAGACCGTTTCCAATAGGAAAGAGTTCGTCAAGCTGTATATTCTCACCCTGTG	120
QY	1916	TTTAACTCTTATGACAG-ACATCATGAGAAACGCTGGCTGGAGAAAGCAAGCTGGA	1974
Db	121	TTTAACTCTTATGACAGTACATCATGAGAAACCTGGACTGGAGAAACCAAGCTGGA	180
QY	1975	ATCAAGATTGCGGGAGAAATAGCAATTAACCTCAGATATGAGATGATACACACCTTATG	2034
Db	181	ATCAAGATTGCTGGAGAAATATCAATTAACCTCAGATATGAGATGATACACACCTTATG	240
QY	2035	GCAGAAAGTAGAGAGAACTAAAAGCCTTGTATGAAGGTGAAAGAGGAGAGTGAATAA	2094
Db	241	GCAGAAAGTAGAGAGAACTAAAAGCCTTGTATGAAGGTGAAAGTGAAGTGAATAA	300
QY	2095	GTTGCTTTAAAGCTCAACATTCAGAAACAGAGATCATGGCATCTGTCCTCCTCACTCA	2154
Db	301	GTTGCTTTAAAGCTCAACATTCAGAAACAGAGATCATGGCATCTGTCCTCCTCACTCA	360
QY	2155	TGGGAAATAGATGGGAAACAGTGGAAACAGTGTGACACTTTATTTTGGGGGGCTCCA	2214
Db	361	TGTTAAATAGATGGGAAACAGTGGAAACAGTGTGACACTTTATTTT-CTGGGCTCCA	418
QY	2215	AAATCAGTCGAGATGCTGACCTGACGCCATGAATTAAGACACTTACTCTCTGGAAGAA	2274
Db	419	AAATCAGTCGAGATGCTGACCTGACGCCATGAATTAAGACACTTACTCTCTGGAAGAA	478
QY	2275	AAGTTA--ACCAACCTAGATAGCATATTGAAAGCAGAGACATTACCTTGGCAACAAGC	2332
Db	479	AAGTTATACCAACCTAGATAGCATATTGAAAGCAGAGACATTACTTTGCCACAAAGG	538
QY	2333	CCCATCTAGTCAAGCTATGTTTTTCCAGTGTGTCATGATGATGATGAGATGGAAGCTG	2392
Db	539	TTTGTCTAGTCAAGCTATGTTTTTCCAGTGTGTCATGATGATGATGAGATGGAAGCTG	598
QY	2393	TGAAGAAAGCTGAGCAGCTGGAAGAAATGATGCTTTTGAACCTGCTGTTGGAGAGACTCT	2452
Db	599	TGAAGAAAGCTGAGCAGCTGGAAGAAATGATGCTTTTGAACCTGCTGTTGGAGAGACTCT	658
QY	2453	TGAGAGTCCCTTGGAGTGCAGAGAGATCCCAACCAAGTCCATTCTGAAGGAGATCAGCCCTG	2512
Db	659	TGAGAGTCCCTTGGAGTGCAGAGAGATCCCAACCAAGTCCATTCTGAAGGAGATCAGCCCTG	718

```

QY 2513 GGATTTCTTTGGAGGAATGATGCTAAAGCTGAACCTCCAGTACTTGGCCACCTGTATCA 2572
Db 719 GGATTTCTTTGGAGGAATGATGCTAAAGCTGAACCTCCAGTACTTGGCCACCTGTATCA 778
QY 2573 GAAGAGCTGACTCAGTGGAAAGACCTGATGCTGGAGGAGATTTGGGGCAGGAGAGAA 2632
Db 779 GAAGAGTTGACCCATTGGAAAGACCTGATGCTGGAGGAGATTTGGGGCAGGAGAGAA 838
QY 2633 GGGGACGACAGAGGATGAGATGCTGGATGGATCACTGACTCGATGGAGCGTGAATCTGG 2692
Db 839 -GGGACAATAGAGGATGAGATGCTGGATGGATCACTGACTCGATGGAGCGTGAATCTGG 897
QY 2693 GTGAATCTCTGGAGTTGGTGATGGACAG 2720
Db 898 GTGAATCTCTGGAGTTGGTGATGGACAG 925

RESULT 22
CC543667/c
LOCUS
DEFINITION
CH240_425B21..T7 CHORI-240 Bos taurus genomic clone CH240_425B21,
genomic survey sequence.
ACCESSION
CC543667
VERSION
CC543667.1 GI:31861951
KEYWORDS
GSS.
SOURCE
Bos taurus (cow)
ORGANISM
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 934)
Holt,R., Stott,J., Yang,G., Barber,S., Smalhus,D., Prabhu,A.-L.,
Teal,M., Cloutier,A., Lee,D., Girm.N., Olson,T., Mayo,M.,
Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R.,
Mathewson,C., Wye,N., Masson,A., Brown-John,M., Jones,S.,
Schein,J., Marra,M., de Jong,P., Keele,J.W. and Kappe,S.M.
Bovine BAC End Sequences from Library CHORI-240, PLATES 399 to 478
Unpublished (2003)
Other GSSs: CH240_425B21.TARBAC13P2
Contact: Rob Holt
Sequencing
The British Columbia Cancer Agency Genome Science Centre
600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
Tel: 604-877-6085
Fax: 604-877-6276
Email: rholt@bccsc.ca
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.choi.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering_information.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
British Columbia Genome Sciences Centre, Canada.
Plate: 425 Row: B Column: 21
Seq primer: 17
Class: BAC ends.
FEATURES
Location/Qualifiers
1..934
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="bred: Hereford"
/db_xref="taxon:9913"
/clone="CH240_425B21"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
Hereford bull Li Domino 99375; CHORI-240 Bovine BAC
library (Male) produced by Pieter de Jong"

Query Match 22.0%; Score 784.2; DB 1; Length 934;

```

```

Best Local Similarity 93.0%; Pred. No. 0.16;
Matches 873; Conservative 0; Mismatches 58; Indels 8; Gaps 5;
QY 1812 TGTATGAGGTCAGAGAGCAACAGTCTAGAACTGGACATGGAACAACAGACTGGTTCAAG 1871
Db 934 TGTATGAGGTCAGAGAGCAACAGTCTAGAACTGGACATGGAACAACAGACT-GTTTCAA 876
QY 1872 TAGAAAGAGAGTATGTCAGGCTGTATATGTCAACGGCTGTTTAACTTCTATGCGAG 1931
Db 875 TAGAAAGAGAGTATGTCAGGCTGTATATGTCAACGGCTGTTTAACTTCTATGCGAG 816
QY 1932 AG-ACATCATGAGAAACGCTGGGTGGAAGACACAAGCTGGAAATCAAGATTCCCGGGA 1990
Db 815 AGTACATTATGAGAAACGCTGGGTGGAAGACACAAGCTGGAAATCAAGATTCCCGGGA 756
QY 1991 GAAATAGCAATAACCTCAGATATATGACATGATACACCCCTTATGGCAGAAAGTGAAGG 2050
Db 755 GAAATATCAATAACCTCAGATATATGACATGATACACCCCTTATGGCAGAAAGTGAAGG 696
QY 2051 AACTTAAAGCTCTTGTAGTGAAGTGAAGAGGAGAGTGAAGAAAGTTGGCTTAAAGCTCA 2110
Db 695 AACTTAAAGCTCTTGTAGTGAAGTGAAGAGGAGAGTGAAGAAAGTTGGCTTAAAGCTCA 636
QY 2111 ACATTTCAGAAAACGAGATCATGTCATCTGGTCCCATCACTTCATGGGAAATAGATGGG 2170
Db 635 ACATTTCAGAAAACGAGATCATGTCATCTGGTCCCATCACTTCATGGGAAATAGATGGG 576
QY 2171 AAACAGTGAACACAGTGTCAAGCTTTATTTTGGGGGGCTCCAAATCAGTCGAGATGG 2230
Db 575 AAACAGTGAACACAGTGTCAAGCTTTATTTTCTGGGCTCCAAATCAGTCGAGATGG 518
QY 2231 TGACTGCGACCATGAATTTAAAGACACTTACTCTCTGGAGAAAGTTA--ACCAACT 2288
Db 517 TGATTGCGAGCAATGAATTTAAAGATGCTTACTCTCTGGAGAAAGTTATGACCACT 458
QY 2289 AGATAGCATATTGAAAAGCAGACAGATTAACCTTGCCCAACAAAGCCCATCTAGTCAAGC 2348
Db 457 AGATAGCATATTGAAAAGTGAAGACATTAACCTTGCCCAACAAAGTTCATCTAGTCAAGC 398
QY 2349 TATGTTTTTCCAGTGGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2408
Db 397 TATGTTTTTCCAGTGGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 338
QY 2409 CTGAAGATTGATGCTTTTGAACCTGTGGTGGTGGAGAGACTCTTGGAGTCCCTTGGAC 2468
Db 337 CCAGAGATTGATGCTTTTGAACCTGTGGTGGTGGAG--ACTCTTGGAGTCCCTTGGAC 280
QY 2469 TGAAGGAGATCCACCAAGTCCATCTGAAGAGATCAGCCCTGGGATTTCTTTGGAGG 2528
Db 279 TGAAGGAGATCCACCAAGTCCATCTGAAGAGATCAGCCCTGGGATTTCTTTGGAGG 220
QY 2529 AATGATGCTAAAGCTGAAACTCCAGTACTTTGGCCACCTGATCAGAGAGCTGACTCACT 2588
Db 219 AATGATGCTAAAGCTGAAACTCCAGTACTTTGGCCACCTCATGCGAAGAGTTGACTCAT 160
QY 2589 GGAAGAGACCTGATCTGGAGGGATTTGGGGCAGGAGAGAGAGGGGACGACAGAGAT 2648
Db 159 GGAAGAGACCTGATCTGGAGGGATTTGGGGCAGGAGAGAGGGGATGACAGAGAT 100
QY 2649 GAGATGCTGGATGGCATCACTGACTGATGAGCTGCTGGTGAACCTCTCTGGAGTT 2708
Db 99 GAGATGCTGGAGGGCATCACTGACTGATGAGCTGATGAGTCTGAGTGAATCCAGAGAT 40
QY 2709 GGTGATGGACAGGAGGCGCTGTCTCGGCGCATTCATGG 2747
Db 39 GGTGATGGACAGGAGGCGCTGTCTCGGCGCATTCATGG 1

```

```

RESULT 23
CC547247
LOCUS
DEFINITION
CH240_430B11..TARBAC13P2 CHORI-240 Bos taurus genomic clone
CH240_430B11, genomic survey sequence.

```

```

ACCESSION  CC547247
VERSION     CC547247.1  GI:31865531
KEYWORDS   GSS
SOURCE     Bos taurus (cow)
ORGANISM   Bos taurus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
           Bovidae; Bovinae; Bos.
REFERENCE  1 (bases 1 to 907)
AUTHORS   Holt, R., Stott, J., Yang, G., Barber, S., Smailus, D., Prabhu, A.-L.,
           Tsai, M., Cloutier, A., Lee, D., Girm, N., Olson, T., Mayo, M.,
           Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
           Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
           Schein, J., Marra, M., de Jong, P., Keele, J.W. and Kappes, S.M.
           Bovine BAC End Sequences from Library CHORI-240, PLATES 399 to 478
           Unpublished (2003)
TITLE     Bovine BAC End Sequences from Library CHORI-240, PLATES 399 to 478
JOURNAL
COMMENT   Other GSSs: CH240_430B11.77
           Contact: Rob Holt
           Sequencing
           The British Columbia Cancer Agency Genome Science Centre
           600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
           Tel: 604-877-6085
           Fax: 604-877-6276
           Email: rholt@bccsc.ca
           Clones are derived from the bovine BAC library CHORI-240
           (http://www.chori.org/bacpac/bovine240.htm). For BAC library
           availability, please contact Pieter de Jong (pdejong@mail.cho.org).
           Clones may be purchased from BACPAC Resources
           (http://www.chori.org/bacpac/ordering/information.htm). This work
           was undertaken as part of the International Bovine BAC Mapping
           Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
           British Columbia Genome Science Centre, Canada.
           Plate: 430 row: B column: 11
           Seq primer: SP6
           Class: BAC ends.
FEATURES   Location/Qualifiers
           1..907
           /organism="Bos taurus"
           /mol_type="genomic DNA"
           /strain="breed: Hereford"
           /db_xref="taxon:9913"
           /clone="CH240_430B11"
           /sex="Male"
           /cell_type="Blood"
           /clone_lib="CHORI-240"
           /note="vector: pFARBAC1.3; Site 1: MboI; Site 2: MboI;
           Hereford bull Li Domino 99375; CHORI-240 Bovine BAC
           library (Male) produced by Pieter de Jong"
Query Match 21.98; Score 783.8; DB 1; Length 907;
Best Local Similarity 94.18; Pred. No. 0.16;
Matches 855; Conservative 0; Mismatches 48; Indels 6; Gaps 4;
QY 1796 TGACTCTGAAAATTTGTATGACGTCAGGAGCAACACAGTTAGAACTGGACATGGAACA 1855
Db 1 TGCTCTTGAGAAATTTGTATGACGTCAGGAGCAACACAGTTAGAACTGGACATGGAACA 60
QY 1856 ACAGACTGGTTCCAGTAGGAAAAGAGATGTCACAGGCTGTATATGTACCCGGCTTG 1915
Db 61 ACAGACTGGTTCCAGTAGGAAAAGAGATGTCACAGGCTGTATATGTACCCCTGTTTA 120
QY 1916 TTTAACTTCTATGCAGAG-ACATCATGAGAAACCTTGGGCTGGAAGACACAAAGCTGGA 1974
Db 121 TTTAACTTATATGCAGTAGTACATCATGAGAAAGC-GGACTGGAAGAAACACAACTGGA 179
QY 1975 ATCAAGATTGCCCGGAGAAATACCAATACCTCAGATATGCAGATGATACCACTTATG 2034
Db 180 ATCAAGATTGCCCGGAGAAATATCAATCACTCAGATATGCAGATGATACCACTTATG 239
QY 2035 GCAGAAAGTGAAGAGGAACCTTAAAGCCCTTTGATCAAGGTGAAGAGGAGAGTGA AAAA 2094
Db 240 GCAGAAAGTGAAGAGGAACCTTAAAGCCCTTTGATCAAGGTGAAGAGTGAAGAGTGA AAAA 299
2095 GTTGGCTTAAAGCTCAACATTCAGAAAACGAAGATCATGGCATCTGGTCCCATCACTTCA 2154
300 GTTGGCTTAAAGCTCAACATTCAGAAAACGAAGATCATGGCATCTGGTCCCATCACTTCA 359
2155 TGGGAATAGATGGGAAAACAGTGGAAAACAGTGTGACAGCTTTTATTTTGGGGGCTCCA 2214
360 TGGGAATAGATGGGAAAACAGTGGAAAACAGTGTGACAGCTTTTATTTT--CTGGGCTCCA 417
2215 AATCATCTCAGATGGTGTGATCGACGCATGAAATTAAGACACACTTACTCTTGGAGAA 2274
418 AATCATCTCAGATGGTGTGATCGACGCATGAAATTAAGACACACTTACTCTTGGAGAA 477
2275 AAGTTA--ACCAACCTTAGATAGCATATTGAAAGCAGACAGACATTACCTTGCACAAAGC 2332
478 AAGTTATGACCAACCTTAGATAGCATATTGAAAGCAGACAGACATTACTTTGCCAACAAAGG 537
2333 CCATCTAGTCAAGGCTATGGTTTTTCCAGTGGTGCATGTATGATGTGAGAGTTGAGCTG 2392
538 TTTCGCTAGTCAAGGCTATGGTTTTTCTGTGGTCAATGTATGATGTGAGAGTTGAGCTG 597
2393 TGAAGAAAGCTCAGCACTGAGAAATGATGCTTTTGAAGTGTGGTGTGGAGAAAGACTCT 2452
598 TGAAGAAAGCTCAGCGCCAAATAATGATGCTTTTGAAGTGTGGTGTGGAGAAAGACTCT 657
2453 TGAGAGTCCCTTGGACTGCAAGGAGATCCAAACAGTCCATTTCTGAAGGAGATCAGCCCTG 2512
658 TGAGAGTCCCTTGGACTGCAAGGAGATCCAAACAGTCCATTTCTGAAGGAGATCAGCCCTG 717
2513 GAATTTCTTGAAGAAATGATGCTTAAAGCTGAAACTCCAGTACTTTGGCCACTGATCA 2572
718 GGATATCTTTGGAAGGACTGATGCTTAAAGCTGAAACTCCAGTACTTTGGCCACTGATGC 777
2573 GAAGAGCTGACTCACTGCAAAAAGCCCTGATGCTGGGAGGATTTGGGGCAGGAGGAGAA 2632
778 GAAGAGTGTGACTCATTTGAAAAGACTCTGATGCTGNGAGGATTTGGGGCAGGAGGAGAA 837
2633 GGGGACGACAGAGGATGAGTGGTGGATGCATCATCTGATCGATGACATGAGCTGCTGG 2692
838 GGGGACACAGAGGATGAGTGGTGGATGCATCATCTGATCGATGACATGAGCTGAGTCTGA 897
2693 GTGACTCC 2701
898 GTGAACTCC 906
CC480290 907 bp DNA linear GSS 16-JUN-2003
CH240_307N7.T7 CHORI-240 Bos taurus genomic clone CH240_307N7,
genomic survey sequence.
CC480290
CC480290.1 GI:31758678
GSS.
Bos taurus (cow)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
Bovidae; Bovinae; Bos.
1 (bases 1 to 907)
Holt, R., Stott, J., Yang, G., Barber, S., Smailus, D., Prabhu, A.-L.,
Tsai, M., Cloutier, A., Lee, D., Girm, N., Olson, T., Mayo, M.,
Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
Schein, J., Marra, M., de Jong, P., McWilliam, S., Barris, W.,
Dairymple, B.P. and Tellam, R.
Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
Unpublished (2003)
Other GSSs: CH240_307N7.FARBAC13p2
Contact: Rob Holt
Sequencing
The British Columbia Cancer Agency Genome Science Centre
600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
Tel: 604-877-6085

```

Fax: 604-877-6276
 Email: rholt@bcgsc.ca
 Clones are derived from the bovine BAC library CHORI-240
 (http://www.chori.org/bacpac/bovine240.html). For BAC library
 availability, please contact Pieter de Jong (pdejong@mail.choi.org).
 Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/ordering_information.htm). This work
 was undertaken as part of the International Bovine BAC Mapping
 Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
 British Columbia Genome Sciences Centre, Canada.
 Plate: 307 row: N column: 7
 Seq primer: 17
 Class: BAC ends.

FEATURES

Location/Qualifiers
 1. .949
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /strain="breed: Hereford"
 /db_xref="taxon:9913"
 /clone="CH240_307N7"
 /sex="Male"
 /cell_type="Blood"
 /clone_lib="CHORI-240"
 /note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
 Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
 library (Male) produced by Pieter de Jong"

Query Match 21.9%; Score 782.8; DS 1; Length 907;
 Best Local Similarity 93.2%; Pred. No. 0.17; Mismatches 57; Indels 4; Gaps 2;
 Matches 839; Conservative 0;

467 AAGAGATGAGCAAGCAAAAGAAATACCCAGCTGTGGATGTGACGTGGTGAATATAAGCA 526
 Db ATGACATGGAGCAAGCAAAAGAAATACCCAGCTGTGGATGTGACGTGGTGAATATAAGCA 837

527 AGCTCGATCTGTAAAGAGCAATATTCATAGAGCACTGGAACTGGATGTGACGTGGTGAAT 586
 Db AGCTCGATCTGTAAAGAGCAATATTCATAGAGCACTGGAACTGGATGTGACGTGGTGAAT 778

587 AAGGCAAAATGGAGTGTGCAACAGAGATGGCAAGATGCAATGTCAACATTCAGGAA 646
 Db AGGCAAAATGGAGTGTGCAACAGAGATGGCAAGATGCAATGTCAACATTCAGGAA 718

647 TCAGGCAATAAATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGAT 706
 Db TCAGGCAATAAATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGAT 658

707 ACTGGCGGAGCAATCCCTTCAGAGAAATGGAGTAGCCATCATGTGCAACAAAAGAGTCC 766
 Db ACTGGCGGAGCAATCCCTTCAGAGAAATGGAGTAGCCATCATGTGCAACAAAAGAGTCC 538

767 GAAATGCACTATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGAT 826
 Db GAAATGCACTATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGAT 538

827 GCAAAACATTCATATCAGTAATCCAGTCTATGCCCAACAGTAATGCTGCAAGAG 886
 Db GCAAAACATTCATATCAGTAATCCAGTCTATGCCCAACAGTAATGCTGCAAGAG 478

887 CTGAAGTTGAACGGCTCTATAGAGACCTTACAGACCTTTTGAACCTTACACCCAAAAAG 946
 Db CTGAAGTTGAACGGCTCTATAGAGACCTTACAGACCTTTTGAACCTTACACCCAAAAAG 418

947 ATGTCTCTCTATATAGGGGACTGGAATGAAAGTAGGAGCAAGAAACACCTCGAG 1006
 Db ATGTCTCTCTATATAGGGGACTGGAATGAAAGTAGGAGCAAGAAACACCTCGAG 358

1007 TAACAGGCAAAATTTGGCTTGAATACCGAATAGAGAGGCGCAAGCAAGCAATAGAGTTTT 1066
 Db TAACAGGCAAAATTTGGCTTGAATACCGAATAGAGAGGCGCAAGCAAGCAATAGAGTTTT 298

1067 GCCAGAAATCACTGTCTAGCAACACCCCTCTCCACACACAGAGAGAGCTCT 1126

Db 297 GCCAAGAAAATGTACTGTGATAGCAAAACACCCCTCTTCCAAACAACAAGAGACTTT 238
 QY 1127 ACACATGACATCCACGATGCTCAACACACGAAATCAGATTGATTATATCTTTGACGCC 1186
 Db 237 ACACATGACATCCACGATGCTCAACACACGAAATCAGATTGATTATATCTTTGACGCC 178
 QY 1187 AAAGATGGAGAGCTCTATACAGTCAGCAAAACCAAGACAGGAGCTTACTGTGGCTCAG 1246
 Db 177 AAAGATGGAGAGCTCTATACAGTCAGCAAAACCAAGACAGGAGCTTACTGTGGCTCAG 121
 QY 1247 ATCATGAACTCTTATGCCAAATTCAGACTTAATTTGAAGAAAGTAGGAAACCACTA 1306
 Db 120 ACCATGAACTCTTATGCCAAATTCAGACTTAATTTGAAGAAAGTAGGAAACCACTA 61
 QY 1307 GATCACTCAGGTAAGACCTTAATCCATCCCTTATGATTATACAGTGGAAAGTGAATA 1366
 Db 60 AACGATTCAGGATGACCTTAATCAAATCCCTTATGATTATACAGTGAAGAGAAATAGA 1

RESULT 25

CC589137 949 bp DNA linear GSS 18-JUN-2003
 CH240_388A3.T7 CHORI-240 Bos taurus genomic clone CH240_388A3,
 genomic survey sequence.

ACCESSION

CC589137
 CC589137.1 GI:31944510

KEYWORDS

GSS.

SOURCE

Bos taurus (cow)

ORGANISM

Bos taurus

REFERENCE

AUTHORS

Holt, R., Stott, J., Yang, G., Barber, S., Smailus, D., Prabhu, A.-L.,
 Tsai, M., Cloutier, A., Lee, D., Ginn, N., Olson, T., Mayo, M., Chiu, R.,
 Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, A., Chan, A., Chiu, R.,
 Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
 Schein, J., Marra, M., de Jong, P., McWilliam, S., Barris, W.,
 Dalrymple, B.P. and Tellam, R.

TITLE

Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398

JOURNAL

Unpublished (2003)

COMMENT

Other GSSs: CH240_388A3.TARBAC13P2
 Contact: Rob Holt
 Sequencing
 The British Columbia Cancer Agency Genome Science Centre
 600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4S6
 Tel: 604-877-6085
 Fax: 604-877-6276
 Email: rholt@bcgsc.ca

Clones are derived from the bovine BAC library CHORI-240
 (http://www.chori.org/bacpac/bovine240.htm). For BAC library
 availability, please contact Pieter de Jong (pdejong@mail.choi.org).
 Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/ordering_information.htm). This work
 was undertaken as part of the International Bovine BAC Mapping
 Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
 British Columbia Genome Sciences Centre, Canada.

Plate: 388 row: A column: 3
 Seq primer: 17

Class: BAC ends.

FEATURES

Location/Qualifiers

1. .949

/organism="Bos taurus"

/mol_type="genomic DNA"

/strain="breed: Hereford"

/db_xref="taxon:9913"

/clone="CH240_388A3"

/sex="Male"

/cell_type="Blood"

/clone_lib="CHORI-240"

/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
 Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
 library (Male) produced by Pieter de Jong"

Db 221 GAATACAGGCAAAATTTGGCTTGGAGTACGGAATGAAGCAGGCAAAAGGCTTAATAGAGT 162
 QY 1064 TTTGCCAAGAAATGACCTGGTCAATAGCAAAACACCTCTTCCAAACACACAGAGAAC 1123
 Db 161 TTTGCCAAGAACACACCTGGTCAATAGCAAAACACCTCTTCCAAACACAGAGAAC 102
 QY 1124 TCTACATGGACATCACAGATGTCACACCGAAATCAGATTGATTATTTCTTGA 1183
 Db 101 TTTACATGGACATCACAGATGTCACACCGAAATCAGATTGATTATTTCTTGA 42
 QY 1184 GCC-AAAGATGAGAGCTCTATACAGTCAGCAAAACAAAG 1223
 Db 41 GCCGAAAGATGAGAGCTCTATACAGTCAGCAAAACAAAG 1

RESULT 28
 CC511457/c
 LOCUS
 DEFINITION CH240_355A2.T7 CHORI-240 Bos taurus genomic clone CH240_355A2,
 Genomic survey sequence.
 ACCESSION CC511457
 VERSION CC511457.1 GI:31829745
 KEYWORDS GSS.
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus

REFERENCE
 AUTHORS
 Holte, R., Stott, J., Yang, G., Barber, S., Smillus, D., Prabhu, A.-L.,
 Tsai, M., Cloutier, A., Lee, D., Girm, N., Olson, T., Mayo, M.,
 Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
 Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
 Schein, J., Marra, M., de Jong, P., McWilliam, S., Barris, W.,
 Dalrymple, B. P. and Tellam, R.

TITLE Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
 JOURNAL Unpublished (2003)
 COMMENT Other GSSs: CH240_355A2.TARBAC13P2
 Contact: Rob Holte
 Sequencing
 The British Columbia Cancer Agency Genome Science Centre
 600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4S6
 Tel: 604-877-6085
 Fax: 604-877-6276
 Email: rholte@bcgsc.ca

Clones are derived from the bovine BAC library CHORI-240
 (http://www.chori.org/bacpac/bovine240.htm). For BAC library
 availability, please contact Pieter de Jong (pdejong@mail.cho.org).
 Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/ordering/information.htm). This work
 was undertaken as part of the International Bovine Mapping
 Consortium (IBMC) by CSIRO Livestock Industries, Australia and the
 British Columbia Genome Sciences Centre, Canada.
 Plate: 355 row: A column: 2
 Seq primer: T7
 Class: BAC ends.

FEATURES
 source
 1. 945
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /strain="breed: Hereford"
 /db_xref="taxon:9913"
 /clone="CH240_355A2"
 /sex="Male"
 /cell_type="Blood"
 /clone_lib="CHORI-240"
 /note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
 Hereford bull Li Domingo 99375; CHORI-240 Bovine BAC
 library (Male) produced by Pieter de Jong"

Query Match 21.8%; Score 778.1; DB 1; Length 945;
 Best Local Similarity 92.0%; Pred. No. 0.17;

Matches 873; Conservative 0; Mismatches 69; Indels 7; Gaps 5;
 QY 1815 ATCAGGTGAGGAGCAACAGTTAGAACTGGACATGGAACACAGACTGGTTCCTCAAGTAG 1874
 Db 945 ATCAGGTGAGGAGCAACAGTTAGAACTGGACATGGAACACAGACTGGTTCCTCAAGTAG 886
 QY 1875 GAAAGAGGATGCTCAAGGCTGTATTTGTCACCCGGCTTTGTTAACTTCTATGAGAG- 1933
 Db 885 GAAAGAGGATGCTCAAGGCTGTATTTGTCACCCGGCTTTGTTAACTTCTATGAGAGT 826
 QY 1934 ACATCATGAGAAACGCTGGCTGGAGAGACAAAGCTGGAATCAAGATTTGCCGGAGAA 1993
 Db 825 ACATCATGAGAAACGCTGGAGAGAGAAACAAAGCTGGAATCAAGATTTGCCGGAGAA 766
 QY 1994 ATAGCAATTAACCTCAGATATGAGATGATACCAACCTTATGCGAGAAAGTGAAGGAGAC 2053
 Db 765 ATATCAATTAACCTCAGATATGAGATGATACCAACCTTATGCGAGAAAGTGAAGGAGAC 707
 QY 2054 TAAAAAGCCTCTTGTATGAAGTCAAGAGGAGAGTGAAGAAAGTGGCTTTAAAGCTCAACA 2113
 Db 706 TAAAAAGCCTCTTGTATGAAGTCAAGAGTGAAGTGAAGTGAAGTGGCTTTAAAGCTCAACT 647
 QY 2114 TTCAGAAACGAGATCATGGCATCTGGTCCCATCACTTCATGGAAGATAGATGGGAAA 2173
 Db 645 TTCAGAAATGAAGATCATGGCATCTGGTCCCATCACTTCATGGAAGATAGATGGGAAA 587
 QY 2174 CAGTGGAAACAGTGTGAGACTTTTATTTTGGGGGGCTCCAAAATCACTGCAGATGGTGA 2233
 Db 586 CAATAGAAACAGTGTGAGACTTTTATTTT-CTGGGCTCCAAAATCGTACAGATGGTGA 529
 QY 2234 CTGAGCCATGAATTAAGACACTTACTCTCTGGAGAAAGTTA--ACCAACTAGA 2291
 Db 528 CTGAGCCATGAATTAAGAGATGCTTACTCTCTGGAGAAAGTTATGACCAACTAGA 469
 QY 2292 TAGCATATTTGAAAGCAGAGACATTTACCTTGGCAACAAAGCCCATCTAGTCAGAGCTAT 2351
 Db 468 TAGCATATTTCAAGACAGAGACATTTACTTGGCCAAAGTTTCTCTAGTCAGAGCTAT 409
 QY 2352 GGTTCCTCAGTGTGATGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 2411
 Db 408 GGTTCCTCAGTGTGATGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 349
 QY 2412 AAGAATTCATGCTTTTGAACCTGCTGTTGAGAGAGACTCTTGAGAGTCCCTTGAGCTGC 2471
 Db 348 ACGAATTCATGCTTTTGAACCTGCTGTTGAGAGACTCTTGAGAGTCCCTTGAGCTGC 289
 QY 2472 AAGGAGATCCAAACAGTCCATTCTGAAGAGATCAGCCCTGGGATTTCTTTGGAAGGAAT 2531
 Db 288 AAGGAGATCCAAACAGTCCATTCTGAAGAGATCAGCCCTGGGATTTCTTTGGAAGGAAT 229
 QY 2532 GATGCTAAGCTGAACCTCAGTACTTTGGCCACTGATCAGAGAGCTGACTCACTGGA 2591
 Db 228 GATGCTAAGCTGAACCTCAGTACTTTGGCCACTGATCAGAGAGCTGACTCACTGGA 169
 QY 2592 AAAGACCTGATGCTGGAGGGATTTGGGGCAGAGAGAGGGGACGACAGAGATGAG 2651
 Db 168 AAAGACCTGATGCTGGAGGGATTTGGGGCAGAGAGAGGGGACGACAGAGATGAG 110
 QY 2652 ATGCTGATGAGATCAGTACTGATGAGAGCTGAGTCTGGGTGAATCTCTGGAGTTGGT 2711
 Db 109 ATGCTGATGAGATCAGTACTGATGAGAGCTGAGTCTGGGTGAATCTCTGGAGTTGGT 50
 QY 2712 GATGACACAGGAGGCTCTCTCTGGGGATTTCTAGGGGTGACAAAGAT 2760
 Db 49 GATGACACAGGAGGCTCTCTCTGGGGATTTCTAGGGGTGACAAAGTGT 1

RESULT 29
 CC495104
 LOCUS
 DEFINITION CH240_330L9.T7 CHORI-240 Bos taurus genomic clone CH240_330L9,
 genomic survey sequence.
 ACCESSION CC495104
 CC495104 833 bp DNA linear GSS 17-JUN-2003
 CH240_330L9.T7 CHORI-240 Bos taurus genomic clone CH240_330L9,
 genomic survey sequence.

VERSION	CC495104.1	GI:31808185
KEYWORDS	GSS.	
SOURCE	Bos taurus (cow)	
ORGANISM	Bos taurus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.	
AUTHORS	Holt, R., Stott, J., Yang, G., Barber, S., Smailus, D., Prabhu, A.-L., Tsai, M., Cloutier, A., Lee, D., Girm, N., Olson, T., Mayo, M., Butterfield, V., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R., Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S., Schein, J., Warra, M., de Jong, P., McWilliam, S., Barris, W., Dalrymple, B. P. and Tellam, R.	
TITLE	Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398	
JOURNAL	Unpublished (2003)	
COMMENT	Other GSSs: CH240 330L9.TARBAC13P2 Contact: Rob Holt_330L9.TARBAC13P2 Sequencing The British Columbia Cancer Agency Genome Science Centre 600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6 Tel: 604-877-6085 Fax: 604-877-6276 Email: rholt@bcgsc.ca Clones are derived from the bovine BAC library CHORI-240 (http://www.chori.org/bacpac/bovine240.htm). For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/ordering-information.htm). This work was undertaken as part of the International Bovine BAC Mapping Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the British Columbia Genome Sciences Centre, Canada. Plate: 330 row: L column: 9 Seq primer: T7 Class: BAC ends.	
FEATURES	Location/Qualifiers	
source	1..833	
	/organism="Bos taurus"	
	/mol_type="genomic DNA"	
	/strain="bred: Hereford"	
	/db_xref="taxon:9913"	
	/clone="CH240_330L9"	
	/sex="Male"	
	/cell_type="Blood"	
	/clone_lib="CHORI-240"	
	/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI; Hereford bull LI Domino 99375; CHORI-240 Bovine BAC Library (Male) produced by Pieter de Jong"	
Query Match	21.7%; Score 775.4; DB 1; Length 833;	
Best Local Similarity	95.7%; Pred. No. 0.19;	
Matches	797; Conservative 0; Mismatches 36; Indels 0; Gaps 0;	
QY	352	CCCATGACAGTATGAAAGGCAAAATGATAGGATCTCAAGAGGAACTCCCGGTCA 411
DB	1	CCCATGAATAGTATGAAAGGCAAAATGATAGGATCTCAAGAGGAACTCCCGGTCC 60
QY	412	GTAGTGCCCATATGCTACTCGAGATCAGTGCAGAAATACTCCAGAAAGAAATGAAG 471
DB	61	GTAGGGGCCATAGCTACTGAGATCTACTGGAGAAATTAATCTCAGAAAGATGAAGGG 120
QY	472	ATFGAGCCAAAGCAAAAAGAAATACCCAGCTGTGGATGTGACTGTGGATATAGCAAGTTC 531
DB	121	ATFGAGCCAAAGCAAAAAGAAATACCCAGCTGTGGATGTGACTGTGGATATAGCAAGTTC 180
QY	532	CGATCTGTAAAGACCAATATTGCAATAGCAACTGGAATGTGAGTCCATGAATCAAGGC 591
DB	181	CGATCTGTAAAGACCAATATTGCAATAGCAACTGGAATGTGAGTCCATGAATCAAGGC 240
QY	592	AAATTGGAAAGTGGTCAAAACAGAGATGCGAAGAGTGAATGTCAACATTCTAGGAATCAGC 651
DB	241	AAATTGGAAAGTGGTCAAAACAGAGATGCGAAGAGTGAATGTGCAATTATAGGAATCAGC 300
QY	652	GAACTAAATGCACTGGATGGTGAATTTAACTCAGATGCCATTATATCTACTACTGC 711
DB	301	AAACTGAAATGCACTGGATGGTGAATTTAACTCAGATGCCATTATCTACTACTGT 360
QY	712	GGCCAGGAATCCCTCAGAAAGAAATGGAGTAGCCATCATGGTCAACAAAAGAGTCCGAAT 771
DB	361	GGACAGGAATCCCTCAGAAAGAAATGGAGTAGCCATCATGGTCAACAAAAGAGTCCGAAT 420
QY	772	GCAGTACTTGGATGAGTCTCAAAAACACAGAAATGATCTCTGTTGTTTCCAAAGGCAAA 831
DB	421	GCAGTACTTGGATGAGTCTCAAAAACACAGAAATGATCTCTGTTGTTTCCAAAGGCAAA 480
QY	832	CAATTCAATATCACAGTAAATCCAAAGTCTATGCCCCCAACAGTAATGCTGAAGAACTGAA 891
DB	481	CAATTCAATATCACAGTAAATCCAAAGTCTATGCCCCCAACAGTAATGCTGAAGAACTGAA 540
QY	892	GTGAAACGGTCTCTAGAACACCTCAAGACCTTTTAGAATCAACACCCCAAAAAGATGTC 951
DB	541	GTGAAACAGTTCTATAAAGACCTCAAGACCTTTTAGAATCAACACCCCAAAAAGATGTC 600
QY	952	CTTCTCAATTATAGGGGACTGGAAATGCAAAAGTAGGAAGCAAAAGAAACACCTGGAGTAACA 1011
DB	601	CTTTTCATTATAGGGGACTGGAAATGCAAAAGTAGGAAGTCAAGAAACACCTGGAGTAACA 660
QY	1012	GGCAAAATTTGGCTTTGGAATACGGAATGAAGCAGGCGAAAGACTAATAGAGTTTGGCCAA 1071
DB	661	GGCAAAATTTGGCTTCGGAATACAGAAATGAAGCAGGCGAAAGACTAATAGAGTTTGGCCAA 720
QY	1072	GAAATGCACTGGTCTATGCAACACCCCTCTTCCACACACACAGAGAGAGCTCTACACA 1131
DB	721	GAAATGCACTGGTCTATGCAACACCCCTCTTCCACACACACAGAGAGAGATCTACACA 780
QY	1132	TGGACATCACAGATGGTCAACACCCGAAATCAGATTGATTATATTTTTCAG 1184
DB	781	TGGACATCACAGATGGTCAACACCCGAAATCAGATTGATTATATTTTTCAG 833
RESULT 30		
CC539858		
LOCUS	CH240_420A14.T7 CHORI-240 Bos taurus genomic clone CH240_420A14,	GSS 17-JUN-2003
DEFINITION	genomic survey sequence.	
ACCSSION	CC539858	
VERSION	CC539858.1	GI:31858142
KEYWORDS	GSS.	
SOURCE	Bos taurus (cow)	
ORGANISM	Bos taurus	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.	
REFERENCE	1 (bases 1 to 887)	
AUTHORS	Holt, R., Stott, J., Yang, G., Barber, S., Smailus, D., Prabhu, A.-L., Tsai, M., Cloutier, A., Lee, D., Girm, N., Olson, T., Mayo, M., Butterfield, V., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R., Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S., Schein, J., Warra, M., de Jong, P., Keefe, J. W. and Kappes, S. M.	
TITLE	Bovine BAC End Sequences from Library CHORI-240, PLATES 399 to 478	
JOURNAL	Unpublished (2003)	
COMMENT	Other GSSs: CH240_420A14.TARBAC13P2 Contact: Rob Holt_	
	Sequencing The British Columbia Cancer Agency Genome Science Centre 600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6 Tel: 604-877-6085 Fax: 604-877-6276 Email: rholt@bcgsc.ca Clones are derived from the bovine BAC library CHORI-240 (http://www.chori.org/bacpac/bovine240.htm). For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/ordering-information.htm). This work was undertaken as part of the International Bovine BAC Mapping Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the	

British Columbia Genome Sciences Centre, Canada.

Plate: 420 row: A column: 14
Seq primer: T7
Class: BAC ends.

FEATURES

source

Location/Qualifiers

1..887

/organism="Bos taurus"

/mol_type="genomic DNA"

/strain="breed: Hereford"

/db_xref="taxon:9913"

/clone="CH240_420A14"

/sex="Male"

/cell_type="Blood"

/clone_lib="CHORI-240"

/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;

Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC

library (Male) produced by Pieter de Jong"

Query Match 21.7%; Score 774.7; DB 1; Length 887;

Best Local Similarity 94.0%; Pred. No. 0.18; Indels 5; Gaps 3;

Matches 836; Conservative 0; Mismatches 48; Indels 5; Gaps 3;

1670 AAAGCAAGAGAGCTCCAGAAAACATCTATTCTGCTTTATGACTATGCAAAAGCCCTTT 1729

1 AAAGCAAGAGAGTTCAGAAAACATCTATTCTGCTTTATGACTATGCAAAAGCCCTTT 60

1730 GACTGTGGGGTCACAATAAAGTGTGGAAATCTTGAAGGATGGGAATACCGAGCAC 1789

61 GACTGTGGTGTACACAAACAACTGTAGAAAATCTTCAAGAGATGGGAATACCGAGCAC 120

1790 CTGACTGACTCTTGAAAATTTGTATGAGTCTAGGAGCAACAGTTAGAACTGGACAT 1849

121 CTGACTGACTCTTGAGAAATCTGTATGAGTCTAGGAGCAACAGTTAGAACTGGACAT 180

1850 GGAAACAAGAGTGTTCAGATAGAAAAGAGATATCTCAAGCTGTATATTTGACCC 1909

181 GGAACAAGAGAGTGTTCAGATAGAAAAGAGATATCTCAAGCTGTATATTTGAGCC 240

1910 GCTTTGTTTAACTCTATGTCAGAG-ACATCATGAGAAAGCTGGCTGGAGAGACAA 1968

241 TGTATTATTAATCTCTGAGAGTACATCATGAGAAACATCTGGCTGGAGAGACAA 300

1969 GTTGAATCAAGATTGCCGGAGAAATAGCAATACCTCAGATATGAGATGATACACC 2028

301 GCTGAATCAAGATTGCCGGAGAAATATCAATACCTCAGATATGAGATGATACACC 360

2029 CTTATGGCAGAAAGTGAAGAGAACTAAAGCCCTTTGATGAAGTGAAGAGAGAGT 2088

361 CTTATGGCAGAAAGTGAAGAGAACTCAAAAGCCCTTTGATGAAGTGAAGAGAGAGT 420

2089 GAAAAAGTTGGCTTAAAGCTCAACATTCAGAAAACGAGATCATGGATCTGGTCCCATC 2148

421 GAAAAAGTTGGCTTAAAGCTCAACATTCAGAAAACGAGATCATGGATCTGGTCCCATC 480

2149 ACTTCATGGGAAATAGATGGGAAACAGTGAACAGTGTGAGACTTTATTTTGGGG 2208

481 ACTTCATGGGAAATAGATGGGAAACAGTGAACAGTGTGAGACTTTATTTT-CTGG 538

2209 GCTCCAAATCACTGCAGATGTGACTGCAGCCATGAAATTAAGACACTTACTCTCTG 2268

539 GCTCCAAATCACTGCAGATGTGACTGCAGCCATGAAATTAAGACACTTACTCTCTG 598

2269 GAAGAAAAGTTA-ACCACCTAGATAGCATATTGAAAAGCAGAGACATTACCTTGCAA 2326

599 GAAGAAAAGTTAATGTCACACCTAGATACATATTCAAAGCAGAGACATTACTTTGCCA 658

2327 CAAGCCCCATCTAGTCAAGCTATGTTTTCAGTGGTCAATGATGATGAGAGTT 2386

659 CAAGATCCATCTAGTCAAGCTATGTTTTCAGTGGTCAATGATGATGAGAGTT 718

2387 GGACTGTGAAGAAAGCTGAGCACTGAAGAAATGTATGCTTTTGAAGTGTGTTGGAGAA 2446

719 GGACTGTGAAGAAAGCTGAGCACTGAGCAACCGAAGATTGATGCTTTTGAAGTGTGTTGGAGAA 778

QY 2447 GACTCTTGAGAGTCCCTTGGACTGCAAGGAGATCCCAACAGTCCATCTTGAAGGATCA 2506

Db 779 GACTCTTGAGAGTCCCTTGGACTGCAAGGAGATCCCAACAGTCCATCTTGAAGGATCA 838

QY 2507 GCCTGGGATTTCTTTGGAAGGAATGATGTAAGGTGAAACTCCAGTA 2555

Db 839 GCCTGGGATTTCTTTGGAAGGAATGATGTAAGGTGAAACTCCAGTA 887

RESULT 31

CC528566/c

LOCUS

DEFINITION

CH240_404A15.TARBAC13P2 CHORI-240 Bos taurus genomic clone

CC528566

VERSION

CC528566.1 GI:31846854

KEYWORDS

SOURCE

Bos taurus (cow)

ORGANISM

REFERENCE

AUTHORS

1. (bases 1 to 896)

Holt,R., Stott,J., Yang,G., Barber,S., Smailus,D., Prabhu,A.-L.,

Tsai,M., Cloutier,A., Lee,D., Girn,N., Olson,T., Mayo,M.,

Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R.,

Mathewson,C., Wye,N., Masson,A., Brown-John,M., Jones,S.,

Schein,J., Marra,M., de Jong,P., Keele,J.W. and Kappes,S.M.

Bovine BAC End Sequences from Library CHORI-240, PLATES 399 to 478

Unpublished (2003)

Other_GSSs: CH240_404A15.T7

CONTACT: Rob Holt

Sequencing

The British Columbia Cancer Agency Genome Science Centre

600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6

Tel: 604-877-6085

Fax: 604-877-6276

Email: rholt@bccgsc.ca

Clones are derived from the bovine BAC library CHORI-240

(http://www.chori.org/bacpac/bovine240.htm). For BAC library

availability, please contact Pieter de Jong (pdejong@mail.cho.org).

Clones may be purchased from BACPAC Resources

(http://www.chori.org/bacpac/ordering/information.htm). This work

was undertaken as part of the International Bovine BAC Mapping

Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the

British Columbia Genome Sciences Centre, Canada.

Plate: 404 row: A column: 15

Seq primer: SP6

Class: BAC ends.

FEATURES

source

Location/Qualifiers

1..896

/organism="Bos taurus"

/mol_type="genomic DNA"

/strain="breed: Hereford"

/db_xref="taxon:9913"

/clone="CH240_404A15"

/sex="Male"

/cell_type="Blood"

/clone_lib="CHORI-240"

/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;

Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC

library (Male) produced by Pieter de Jong"

Query Match 21.7%; Score 774.3; DB 1; Length 896;

Best Local Similarity 94.0%; Pred. No. 0.18; Indels 7; Gaps 4;

Matches 846; Conservative 0; Mismatches 47; Indels 7; Gaps 4;

QY 1888 TCAAGCTGTATATTGTCAACCGGCTTTTAACTTCTATGAGAG-ACATCATGAGAAA 1946

Db 896 TCAAGCTGTATATTGTCAACCGGCTTTTAACTTCTATGAGAGTACATCATGAGAAA 837

QY 1947 CGCTGGCTGGAAGAGCAACAGTGGATCAAGATTCCCGGAGAAATAGCAATAACCT 2006

Db 836 CCCTGACTGGAGAAACACACAGCTGGAAATCAGATTCGAGGAAAT--CAATCACT 779
 QY 2007 CAGATATGAGATGATACACCCCTTATGGCAGAAAGTGAAGAGAACTTAAAGACCTCTT 2066
 Db 778 CAGATATGAGATGATACACCCCTTATGGCAGAAAGTGAAGAGAACTTAAAGACCTCTT 719
 QY 2067 GATGAAGGTGAAGAGGAGAGTGAAGAGTGGCTTAAAGCTCAACATTCAGAAAGCGAA 2126
 Db 718 GATGAAGGTGAAGAGGAGAGTGAAGAGTGGCTTAAAGCTCAACATTCAGAAAGCGAA 659
 QY 2127 GATCATGGCATCTGTGCCCATCTCATGCGAAATAGATGGGAAACAGTGGAAACAGT 2186
 Db 658 GATCATGGCATCTGTGCCCATCTCATGCGAAATAGATGGGAAACAGTGGAAACAGT 599
 QY 2187 GTCAGACTTTATTTTGGGGGGCTCCAAATCACTGCAGATGTGATGCTGAGCCATGAA 2246
 Db 598 GTCAGACTTTATTTT--CTGGGCTCCAAATCACTGCAGATGTGATGCTGAGCCATGAA 541
 QY 2247 ATTAAGACACTTACTCTCTGGAAGAAAGTTA--ACCACTTAGATAGCATATTGAAA 2304
 Db 540 ATTAAGACCGTTACTCTCTGGAAGAAAGTTATGACCACTAGTAGCATATTAAAA 481
 QY 2305 AGCAGAGACATTACCTTGCACAAAGCCCATCTAGTCAAGGCTATGTTTTCCAGTG 2364
 Db 480 AGCAGAGACATTACTTTGCCAACAGAGTCCGTCTAGTCAAGGCTATGTTTTCCAGTG 421
 QY 2365 GTCATGATGATGTGAGAGTTGAGCTGTGAAGAAAGCTGAGCACTGAAGAAATTTGATGCT 2424
 Db 420 GTCATGATGATGTGAGAGTTGAGCTGTGAAGAAAGCTGAGCACTGAAGAAATTTGATGCT 361
 QY 2425 TTTGAACTGTGTTTGGAGAGACTCTTGAAGAGTCCCTTGAAGTGGTGAAGAGATCCAAC 2484
 Db 360 TTTGAACTGTGTTTGGAGAGACTCTTGAAGAGTCCCTTGAAGTGGTGAAGAGATCCAAC 301
 QY 2485 CAGTCCATCTCAAGAGATCAGCCCTGGGATTTCTTTGGAAGGAATGATCTAAAGCTG 2544
 Db 300 CAGTCCATCTCAAGAGATCAGCCCTGGGATTTCTTTGGAAGGAATGATCTAAAGCTG 241
 QY 2545 AATCCAGTACTTTGGCCACTGATAGAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 2604
 Db 240 AATCCAGTACTTTGGCCACTGATAGAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 181
 QY 2605 CTGGAGGAGTGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2664
 Db 180 CTGGAGGAGTGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 121
 QY 2665 ATCACTGACTGATGAGCTGAGTCTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2724
 Db 120 ATCACTGACTGATGAGCTGAGTCTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 61
 QY 2725 GCCTGCTCTGGGCGATTTCATGGGCTCACAAGAGTGGACACGACTGAGCACTGAACT 2784
 Db 60 GCCTGCTGCTGGGCGATTTCATGGGCTCACAAGAGTGGACACGACTGAGCACTGAACTGATCT 1

 RESULT 32
 CC508740
 LOCUS
 DEFINITION
 CH240_351A11.TARBAC13P2 CHORI-240 Bos taurus genomic clone
 CH240_351A11, genomic survey sequence.
 ACCESSION
 CC508740
 VERSION
 GSS.
 SOURCE
 Bos taurus (cow)
 ORGANISM
 Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 827)
 REFERENCE
 AUTHORS
 Holt, R., Stott, J., Yang, G., Barber, S., Smalhus, D., Prabhu, A.-L.,
 Tsai, M., Cloutier, A., Lee, D., Girn, N., Olson, T., Mayo, M.,
 Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,

Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
 Schein, J., Marra, M., de Jong, P., McWilliam, S., Barris, W.,
 Dalrymple, B. P. and Tellam, R.
 Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
 Unpublished (2003)
 Other GSSs: CH240_351A11.T7
 Contact: Rob Holt
 Sequencing
 The British Columbia Cancer Agency Genome Science Centre
 600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4S6
 Tel: 604-877-6085
 Fax: 604-877-6276
 Email: rholt@bcgsc.ca
 Clones are derived from the bovine BAC library CHORI-240
 (http://www.chori.org/bacpac/bovine240.htm). For BAC library
 availability, please contact Pieter de Jong (pdejong@mail.cho.org).
 Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/ordering_information.htm). This work
 was undertaken as part of the International Bovine BAC Mapping
 Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
 British Columbia Genome Sciences Centre, Canada.
 Plate: 351 row: A column: 11
 Seq primer: SP6
 Class: BAC ends.
 Location/Qualifiers
 1..827
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /strain="breed: Hereford"
 /db_xref="taxon:9913"
 /clone="CH240_351A11"
 /sex="Male"
 /cell_type="Blood"
 /clone_lib="CHORI-240"
 /note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
 Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
 library (Male) produced by Pieter de Jong"

Query Match 21.7%; Score 174.2; DB 1; Length 827;
 Best Local Similarity 96.0%; Pred. No. 0.2;
 Matches 794; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

 QY 438 TCAGTGGAGAAATAAATCCAGAAAGATGAAGAGATGGAGCCAAAGCAAAAGAAATACCC 497
 Db 1 TCAGTGGAGAAATAAATCCAGAAAGATGTGTTGGATGGAGCCAAAGCAAAAGCAATACCC 60
 QY 498 AGCTGTGATGTGCTGTGATATAGCAAGTCCGATGCTGTAAGAGCAATATTCAT 557
 Db 61 AGCTGTGATGTGCTGTGATAGCAAGTCCGATGCTGTAAGAGCAATATTCAT 120
 QY 558 AGGAACCTGGAATGTGAGTCCATGAATCAAGGCAAAATGGAAGTGTCAAAACAAAGAT 617
 Db 121 AGGAACCTGGAATGTGAGTCCATGAATCAAGGCAAAATGGAAGTGTCAAAACAAAGAT 180
 QY 618 GCAAGAGTGAATGTCAACATTCATAGGAATCAGGAATCAAAATGGAATGGATGGGTGA 677
 Db 181 GCAAGAGTGAATGTGCAATTCATAGGAATCAGGAATCAAAATGGAATGGATGGGTGA 240
 QY 678 ATTTAACTCAGATGACCAATATATCTACTACTCTACTCTGTGGCAGGAATCCCTCAGAAGAAATGG 737
 Db 241 ATTTAACTCAGATGACCAATATATCTACTACTCTGTGGCAGGAATCCCTCAGAAGAAATGG 300
 QY 738 AGTAGCCATGTGTCACAAAGAGTCCGAAATGCAAGTGTGATGCTGAGTGTGAGTGTCAAAAA 797
 Db 301 AGTAGCCATGTGTCACAAAGAGTCCGAAATGCAAGTGTGATGCTGAGTGTGAGTGTCAAAAA 360
 QY 798 CGACAGATGATCTCTGTTTGTTCCTCAAGGCAAAACCAATTCATATCACAGTAATCCAAAGT 857
 Db 361 CGACAGATGATCTCTGTTTGTTCCTCAAGGCAAAACCAATTCATATCACAGTAATCCAAAGT 420
 QY 858 CTATGCCCCCAACAGTAATGTCTGAAGAGCTGAAGTGTGACCGTCTTATGAAGACCTACA 917
 Db 421 CTATGCCCCCAACAGTAATGTCTGAAGAGCTGAAGTGTGATGCTGATGAGACCTACA 480

QY 918 AGACCTTTTAGAAGTAAACCCCAAAAGATGCTCTCTCATTTATAGGGAGTGAATGC 977
 |||||
 Db 481 AGACCTTTTAGAAGTAAATACCCCAAAAGATGCTCTCTCATTTATAGGGAGTGAATGC 540
 |||||
 QY 978 AAAAGTAGGAAGCAAGAAACACCTGGAGTAAACAGGCAAAATTTGGCCCTTGGNAATACGGAA 1037
 |||||
 Db 541 AAAAGTAGGAAGTCAAGAAACACCTGGAGTAAACAGGCAAAATTTGGCCCTTGGNAATACGGAA 600
 |||||
 QY 1038 TGAACAGGGCAAGACATTAAGATTTTGGCAAGAAATGCACTGGTGCATAGCAACAC 1097
 |||||
 Db 601 TGAACAGGGCAAGAGCTTAATAGATTTTGGCAAGAAATGCACTGGTGCATAGCAACAC 660
 |||||
 QY 1098 CCTCTTCCAAACACAAAGAGAGACTCTACACATGGACATCACACAGATGCTCAACACCG 1157
 |||||
 Db 661 CCTCTTCCAAACACATAGAGAGACTCTACACATGGACATCACACAGATGCTCAACACCA 720
 |||||
 QY 1158 AAATCAGATTAATATATCTTTTGCAGCAAGATGGAGAGCTCTATACAGTCAAGCAAA 1217
 |||||
 Db 721 AAATCAGACTGATTATATCTTTTGCAGCAAGATGGAGAGCTCTATACAGTCAAGCAAA 780
 |||||
 QY 1218 AACAGACCAAGAGCTTACTGTGCTCAGATCATGAACTCCTTATTG 1264
 |||||
 Db 781 AACAGACCAAGAGCTGACTGTGGCTCAGATCATGAACTCCTTATTG 827
 |||||

RESULT 33

CC512039

LOCUS

DEFINITION CH240_356E9.T7 CHORI-240 Bos taurus genomic clone CH240_356E9,

genomic survey sequence.

ACCESSION

CC512039

VERSION

CC512039.1

GI:31830327

GSS.

SOURCE

Bos taurus

ORGANISM

Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovidae; Bovinae; Bos.

1 (bases 1 to 898)

Holt, R., Stott, J., Yang, G., Barber, S., Smailus, D., Prabhu, A.-L.,

Tsai, M., Cloutier, A., Lee, D., Girm, N., Olson, T., Mayo, M.,

Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,

Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,

Schein, J., Marra, M., de Jong, P., McWilliam, S., Barris, W.,

Dalrymple, B. P. and Tellam, R.

Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398

Unpublished (2003)

Other GSSs: CH240_356E9.TARBAC13P2

Contact: Rob Holt

Sequencing

The British Columbia Cancer Agency Genome Science Centre

600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6

Tel: 604-877-6085

Fax: 604-877-6276

Email: rholt@bcgsc.ca

Clones are derived from the bovine BAC library CHORI-240

(http://www.chori.org/bacpac/bovine240.htm). For BAC library

availability, please contact Pieter de Jong (pdjong@mail.choi.org).

Clones may be purchased from BACPAC Resources

(http://www.chori.org/bacpac/ordering/information.htm). This work

was undertaken as part of the International Bovine BAC Mapping

Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the

British Columbia Genome Sciences Centre, Canada.

Plate: 356 row: E column: 9

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

1..898

/organism="Bos taurus"

/mol_type="genomic DNA"

/strain="breed: Hereford"

/db_xref="taxon:9913"

/clone="CH240_356E9"

/sex="Male"

/cell_type="Blood"

/clone_lib="CHORI-240"

/notes="Vector: pTARBAC1.3; Site 1: Mbol; Site 2: Mbol;

Hereford bull Li Domino 99375; CHORI-240 Bovine BAC

Library (Male) produced by Pieter de Jong"

Query Match 21.7%; Score 774.2; DB 1; Length 898;
 Best Local Similarity 94.0%; P-red. No. 0.18;
 Matches 846; Conservative 0; Mismatches 48; Indels 6; Gaps 4;

QY 1796 TGACTCTTGAAGAAATTTGTATGCAAGTCAAGGCTGTATATGTGACCCGCTTG 1915
 |||||
 Db 1 TGCTCTTGAAGAAATTTCTATGCAAGTCAAGGCTGTATATGTGACCCGCTTTA 120
 |||||
 QY 1856 ACAGACTGGTTCACAGTAGGAAAGAGTATGTCAAGGCTGTATATGTGACCCGCTTG 1915
 |||||
 Db 61 ACAGACTGGTTCACAGTAGGAAAGAGTATGTCAAGGCTGTATATGTGACCCGCTTTA 120
 |||||
 QY 1916 TTTAACTTCTATGCAAG-ACATCATAGAAACGCTGGCTGGAAGACCAAGCTGGA 1974
 |||||
 Db 121 TTTAACTTCTATGCAAGTACATCATGACACACGCTGGACTGGAAGAAACCAAGCTGGA 180
 |||||
 QY 1975 ATCAAGATTGCCGGAGAAATAGCAATACCTCAGATATGCAAGTACCAACCCCTTATG 2034
 |||||
 Db 181 ATCAAGATTGCCAGAGAAATATCAATACCTCAGATATGCAAGTACCAACCCCTTATG 240
 |||||
 QY 2035 GCAGAAAGTGAAGAGGAACTAAAAAGCCTCTTGTGAAAGGTGAAAGAGGAGAGTGAAGAA 2094
 |||||
 Db 241 GCAGAAAGTGAAGAGGAACTAAAAAGCCTCTTGTGAAAGTGAAGTGGAGAGTGAAGAA 300
 |||||
 QY 2095 GTTGCTTTAAAGCTCAACATTCAGAAACAGAGATCAAGGATCTGGTCCCATCACTTCA 2154
 |||||
 Db 301 ATTGCTTTAAAGCTCAACATACAGAAACAAAGATCATGGCATCCGTTCTTCACTTCA 360
 |||||
 QY 2155 TGGGAAATAGATGGGAAACACAGTGTGAAACAGTGTGCAAGTGTATTTTGGGGGCTCCA 2214
 |||||
 Db 361 TGGGAAATAGATGGGAAACACAGTGTGAAACAGTGTGCAAGTGTATTTT-CTGGGCTCCA 418
 |||||
 QY 2215 AAATCACTGCAAGTGGTGTGCTGCAAGCCATGAAATTAAGACATCTACTCTTGGAGAA 2274
 |||||
 Db 419 AAATCACTGCAAGTGGTGTGCTGCAAGCCATGAAATTAAGACATCTACTCTTGGAGAA 478
 |||||
 QY 2275 AAGT--TAACCAACCTAGATAGCATATTGAAAGCAGAGACATACCTTGTCCCAACAAGC 2332
 |||||
 Db 479 AAGTTATTAACCAACCTAGATAGCATATTCAAAAGCAGAGACATTAATTTGCCAACAAGG 538
 |||||
 QY 2333 CCATCTAGTCAAGGCTATGTTTTTCCAGTGGTCAATGTATGGATGTGAGAGTTGGAGCTG 2392
 |||||
 Db 539 TTCATCTAGTCAAGGCTATGTTTTTCCAGTGGTCAATGTATGGATGTGAGAGTTGGAGCTG 598
 |||||
 QY 2393 TGAAGAAGCTGAGCAGCTGAAGAAATGATGCTTTTGAAGTGTGGTGTGGAGAGATCT 2452
 |||||
 Db 599 TGAAGAAGCTGAGCAGCTGAAGAAATGATGCTTTTGAAGTGTGGTGTGGAGAGATCT 658
 |||||
 QY 2453 TGAGAGTCCCTTGGAGTGAAGAGATCCCAACAGTCCATTTCTGAAGAGATCAGCCCTG 2512
 |||||
 Db 659 TGAGAGTCCCTTGGAGTGAAGAGATCCCAACAGTCCATTTCTGAAGAGATCAGCCCTG 718
 |||||
 QY 2513 GGATTTCTTTGGAAGGAATGATGCTAAAGCTGAAGATCCAGTACTTTGGCCACCTGATCA 2572
 |||||
 Db 719 GGATTTCTTTGGAAGGAATGATGCTAAAGCTGAAGATCCAGTACTTTGGCCACCTCATGC 778
 |||||
 QY 2573 GAAGAGTGTACTCACTGGAAGAAAGACCTGTGCTGGAGGAGATTGGGGGAGGA-GGAGA 2631
 |||||
 Db 779 GAAGAGTGTACTCACTGGAAGAAAGACTCTGTGCTGGAGGAGATTGGGGGAGGAGGA 838
 |||||
 QY 2632 AGGGGAGCAGAGAGATGAGATGGCTGGATGGCATCACTGATCGATGAGAGTGGAGTCTG 2691
 |||||
 Db 839 AGGGGAGCAGAGAGATGAGATGGCTGGATGGCATCACTGATCGATGAGAGTGGAGTCTG 898
 |||||

FEATURES

source

CC581072	881 bp	DNA	linear	GSS 18-JUN-2003					
LOCUS	CH240_376A7.T7	CHORI-240	Bos taurus	genomic clone	CH240_376A7,				
DEFINITION	genomic survey sequence.								
ACCESSION	CC581072								
VERSION	CC581072.1	GI:31928363							
KEYWORDS	GSS.								
SOURCE	Bos taurus (cow)								
ORGANISM	Bos taurus	<p>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.</p> <p>1 (bases 1 to 881) Holt,R., Scott,J., Yang,G., Barber,S., Smailus,D., Prabhu,A.-L., Tsai,M., Cloutier,A., Lee,D., Girm,N., Olson,T., Mayo,M., Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R., Mathewson,C., Wye,N., Masson,A., Brown-John,M., Jones,S., Schein,J., Marra,M., de Jong,P., McWilliam,S., Barris,W., Dalrymple,B.P. and Tellam,R.</p> <p>Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398</p> <p>Unpublished (2003) Other GSSs: CH240_376A7.TARBAC13P2 Contact: Rob Holt Sequencing The British Columbia Cancer Agency Genome Science Centre 600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4B6 Tel: 604-877-6085 Fax: 604-877-6276 Email: rholt@bcgsc.ca Clones are derived from the bovine BAC library CHORI-240 (http://www.chori.org/bacpac/bovine240.htm). For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/ordering/information.htm). This work was undertaken as part of the International Bovine BAC Mapping Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the British Columbia Genome Sciences Centre, Canada. Plate: 376 row: A column: 7 Seq primer: T7 Class: BAC ends.</p>							
REFERENCE									
AUTHORS									
TITLE									
JOURNAL									
COMMENT									
FEATURES	Location/Qualifiers								
source	1..881								
	/organism="Bos taurus"								
	/mol_type="genomic DNA"								
	/strain="breed: Hereford"								
	/db_xref="taxon:9913"								
	/clone="CH240_376A7"								
	/sex="Male"								
	/cell_type="Blood"								
	/clone_lib="CHORI-240"								
	/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI; Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC library (Male) produced by Pieter de Jong"								
Query Match	21.7%; Score 773.7; DB 1; Length 881;								
Best Local Similarity	94.5%; Pred. No. 0.19;								
Matches 832; Conservative	0; Mismatches 43; Indels 5; Gaps 3;								
QY	1796	TGACTCTGAAAATTTGATGTCAGGTCAGGAGCAACAGCTTGAAGTGGACATGGAACA	1855						
DB	1	TGCCTCTTGAGAAATTTGATGTCAGGTCAGGAGCAACAGCTTGAAGTGGACATGGAACA	60						
QY	1856	ACAGACTGGTTCGAAGTAGGAAAAAGGAGTAGTGTCAAGGCTGTATTTATGTACCCGGCTTG	1915						
DB	61	ACAGACTGGTTCGAAGTAGGAAAAAGGAGTAGTGTCAAGGCTGTATTTATGTACCCCTGTTTA	120						
QY	1916	TTTAACTTCTATGACAG-ACATCATGAGAAACGCTGGCTGGGAAGACACAAAGCTGGA	1974						
DB	121	TTTAACTTCTATGACAGATCATCATGAGAAACCGCTGACTTGGGAAGACACAAAGCTGGA	180						
QY	1975	ATCAAGATTGCCGGGAGAAATAGCAATAACCTCAGATATGACATGATACCAACCCCTTATG	2034						
DB	181	ATCAAGATTGCCAGGAGAAATATCAATAAAGCTCAGATATGAGATGACACACCCCTTATG	240						
QY	2035	GCAGAAAGTGAAGAGGAAGTAAAAAGGCTCTTGATGAAGGTGAAGAGGAGAGTGAAGAA	2094						
DB	241	GCAGAAAGTGAAGAGGAAGTAAAAAGGCTCTTGATGAAGGTGAAGAGGAGAGTGAAGAA	300						
QY	2095	GTTGGCTTAAAGCTCAACATTCAGAAAAAGAGATCATGGCATCTGGTCCCATCCTCA	2154						
DB	301	GTTGGCTTAAAGCTCAACATTCAGAAAAAGAGATCATGGCATCTGGTCCCATCCTCA	360						
QY	2155	TGGGAATATAGATGGGAAACAGTGGAAACAGTGTGAGACTTTATTTTGGGGGGCTCCA	2214						
DB	361	TGGGAATATAGATGGGAAACAGTGGAAACAGTGTGAGACTTTATTTTCTGGGGCTCCA	418						
QY	2215	AAATCACTGCAGATGGTGAAGTGGAGCCATGAAATTTAAAGACACTTACTCTTGGAGAA	2274						
DB	419	AAATCACTGCAGATGGTGAAGTGGAGCCATGAAATTTAAAGACACTTACTCTTGGAGAA	478						
QY	2275	AAGTTA--ACCAACTAGATAGCATTTTGAAGAGCAGACACTTACTCTTGGAGAAAGC	2332						
DB	479	AAGTTATGACCAACCTAGATAGCATTTTGAAGAGCAGACACTTACTCTTGGAGAAAGC	538						
QY	2333	CCCATCTAGTCAAGGCTATGTTTTTCCAGTGGTCAATGATGATGATGTGAGAGTTGGACTG	2392						
DB	539	TTTGGTCTAGTCAAGGCTATGTTTTTCCAGTGGTCAATGATGATGATGTGAGAGTTGGACTG	598						
QY	2393	TGAAGAAAGCTGAGACACTGAAGAATGATGCTTTTGAAGTGGTGGTGGAGAGACTCT	2452						
DB	599	TAAAGAGGCTGAGCGCGGAAGAATTTGATGCTTTTGAACCTGTGGTGGAGAGAGTCT	658						
QY	2453	TCAGAGTCCCTTGGACTGCAAGGAGATCCAAACAGTCCATTCCTGAAGGAGATCAGCCCTG	2512						
DB	659	TGAGAGTCCCTTGGACTGCAAGGAGATCCAAACAGTCCATTCCTGAAGGAGATCAGCCCTG	718						
QY	2513	GGATTTCTTTGGAAGGAGATGATGCTTAAAGCTGAAACTCCAGTACTTTTGGCCACTCATCA	2572						
DB	719	GGATTTCTTTGGAAGGAGATGATGCTTAAAGCTGAAACTCCAGTACTTTTGGCCACTCATGT	778						
QY	2573	GAAGAGCTGACTCACTGGAAGAGCCCTGATGCTGGAGGAGATGGGGGAGAGAGAGAA	2632						
DB	779	GAAGAGTTGAAATCATTTGGAAGAGACTCTGATGCTGGAGGAGATTTGGGGGAGAGAGAA	838						
QY	2633	GGGGAGCAGACAGAGGATGAGATGGCTGGATGGCATCACTGA	2672						
DB	839	GGGGAGCAGACAGAGGATGGCTGGATGGCATCACTGA	878						
RESULT 35	CC536979/c	915 bp	DNA	linear	GSS 17-JUN-2003				
LOCUS	CH240_416G6.T7	CHORI-240	Bos taurus	genomic clone	CH240_416G6,				
DEFINITION	genomic survey sequence.								
ACCESSION	CC536979								
VERSION	CC536979.1	GI:31855267							
KEYWORDS	GSS.								
SOURCE	Bos taurus (cow)								
ORGANISM	Bos taurus	<p>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.</p> <p>1 (bases 1 to 915) Holt,R., Scott,J., Yang,G., Barber,S., Smailus,D., Prabhu,A.-L., Tsai,M., Cloutier,A., Lee,D., Girm,N., Olson,T., Mayo,M., Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R., Mathewson,C., Wye,N., Masson,A., Brown-John,M., Jones,S., Schein,J., Marra,M., de Jong,P., Keel,J.W. and Kappes,S.M.</p> <p>Bovine BAC End Sequences from Library CHORI-240, PLATES 399 to 478</p> <p>Unpublished (2003) Other GSSs: CH240_416G6.TARBAC13P2 Contact: Rob Holt Sequencing The British Columbia Cancer Agency Genome Science Centre 600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4B6</p>							
REFERENCE									
AUTHORS									
TITLE									
JOURNAL									
COMMENT									

Tel: 604-877-6085
Fax: 604-877-6276

Email: rholt@bcgsc.ca

Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.choi.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/orderinginformation.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
British Columbia Genome Sciences Centre, Canada.

Plate: 416 row: G column: 6

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

FEATURES

source

1..915

/organism="Bos taurus"

/mol_type="genomic DNA"

/strain="breed: Hereford"

/db_xref="taxon:9913"

/clone="CH240_416G6"

/sex="Male"

/cell_type="Blood"

/clone_lib="CHORI-240"

/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;

Hereford bull Li Domino 99375; CHORI-240 Bovine BAC

library (Male) produced by Pieter de Jong"

Query Match

Best Local Similarity 93.8%; Score 771.3; DB 1; Length 915;

Mismatches 834; Conservative 0; Mismatches 52; Indels 3; Gaps 3;

```

QY 712 GGGCAGGATCCCTCAGAGAAATGGAGTAGCCATCATGTGTCAAC-AAAAGATCCGAAA 770
Db 913 GGCAGGAAATCCCTTAGAAGAAATGGAGTAGCCATCATGTGTCAACAAAAGATCCGAAA 854
QY 771 TGCAGTACTTGGATCGAGTCTCAAAAACGACAGATGATCTCTGTGTTTCCAAAGCAA 830
Db 853 TGCAGTACTTGGATCGAGTCTCAAAAACGACAGATGATCTCTGTGTTTCCAAAGCAA 794
QY 831 ACCATTCAATATCAGTAATCCAAAGTCTATGCCCAACCAAGTAATGCTGAAGAAGCTGA 890
Db 793 ACCATTCAATATCAGTAATCCAAAGTCTATGCCCAACCAAGTAATGCTGAAGAAGCTGA 734
QY 891 AGTTGAACGCTCTATCAAGACCTACAGACCTTTTAGACTACACCCCAAAAAGATGT 950
Db 733 AGTTGAATGGTTCTATGAAGACCTACAGACCTTTTAGAACTAACCCCAAAAAGATGT 674
QY 951 CCTTCTCATTATAGGGGACTGGAATGCAAAAGTAGGAAGCAAAACACCTGGAGTAAC 1010
Db 673 CCTTTTCATTATAGGGGACTAGAAATGCAAAAGTAGGAAGCAAAACACCTGGAGTAAC 614
QY 1011 AGGCBAATTTGGCTTTGGAATAGGATGAGAGGCAAGCAAGCAAGCAAGCAAGCAAGCA 1070
Db 613 AGGCBAATTTGGCTTTGGAATAGGATGAGAGGCAAGCAAGCAAGCAAGCAAGCAAGCA 554
QY 1071 AGAAAATGCACTGGTTCATAGCAAAACCCCTCTTCCCAACCAACCAAGCAAGCAAGCA 1130
Db 553 AGAAGACCACTGGTTCATAGCAAAACCCCTCTTCCCAACCAACCAAGCAAGCAAGCA 494
QY 1131 ATGACATCACCAGATGGTCAACACCGGAATCAGATTGATTATTTCTTTGAGGCAAG 1190
Db 493 ATGACATCACCAGATGGTCAACACCGGAATCAGATTGATTATTTCTTTGAGGCAAG 434
QY 1191 ATGAGAGCTCTATACAGTCAGCAAAAACAGACCAAGGAGTCTACTGTGGCTCAGATCA 1250
Db 433 ATGAGAGCTCTATACAGTCTGCAAAAACAGACCAAGGAGTCTACTGTGGCTCAGATCA 374
QY 1251 TGAATCTCTTATTTGCAAAATTCAGCTTAAATTTGAAGAAGTAGGGAAGAAA-CCACTAGAT 1309
Db 373 TGAGCTCTCTTATTTG-CAAAAATCAGACTTAAATTTGAAGAAGTAGGGAAGAAAACCCCTGGAC 315
QY 1310 CACTCAGGTAAGACCTAAATCCATCCCTTATCATTTATACATGAGTGAAGTGAAGATAGAT 1369

```

```

Db 314 CATTCCGGTATGACCTTAATCAANTCCCTTATGATTATACATGGAAGTGACAAATAGAT 255
QY 1370 TTAAGGCCCTAGATCTGTATAGACAGAGTAGTACCTAATGAATGAGTATGGACAGAGTTTCATGACA 1429
Db 254 TTAAGGGACTAGATCTGTATAGATAGAGTGCCTGATGAATGAGTATGGACGGAGGTTTGTGACA 195
QY 1430 TTGTACAGAGACAGAGTGGAGTGGACCATCCCATGGAAGAAATGCAAAAAGCAAAAT 1489
Db 194 TTGTACAGAGACAGAGTGGAGTGGACCATCCCATGGAAGAAATGCAAAAAGCAAAAT 135
QY 1490 GGCTGTCTGGGAGGCTTTACAAATAGCTGTGAAAAGAGAGAAAGTGAAGCAAGGAA 1549
Db 134 GGCTGTCTGGGAGGCTTTACAAATAGCTGTGAAAAGAGAGAGAGGCAAGGCAAGGAG 75
QY 1550 AAAAGGAAAGATATAAGCATCTGAATGAGAGTTCCTCAAGAACTTCCAG 1598
Db 74 AAAAGGAAAGATATAAGCATCTGAATGAGAGTTCCTCAAGAACTTCCAG 25

CC505149 846 bp DNA linear GSS 17-JUN-2003
CH240_345P10.77 CHORI-240 Bos taurus genomic clone CH240_345P10,
genomic survey sequence.
CC505149
CC505149.1 GI:31823442
GSS.
Bos taurus (cow)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 846)
Holt, R., Stott, J., Yang, G., Barber, S., Smal, D., Prabhu, A., L.,
Tsai, M., Cloutier, A., Lee, D., Girm, N., Olson, T., Mayo, M.,
Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
Schein, J., Marra, M., de Jong, P., McWilliam, S., Barris, W.,
Dalrymple, B.P. and Tellam, R.
Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
Unpublished (2003)
Other_GSS: CH240_345P10.TARBAC13P2
Contact: Rob Holt
Sequencing
The British Columbia Cancer Agency Genome Science Centre
500 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4S6
Tel: 604-877-6085
Fax: 604-877-6276
Email: rholt@bcgsc.ca
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.choi.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/orderinginformation.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
British Columbia Genome Sciences Centre, Canada.
Plate: 345 row: P column: 10
Seq primer: T7
Class: BAC ends.
Location/Qualifiers
1..846
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="breed: Hereford"
/db_xref="taxon:9913"
/clone="CH240_345P10"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
Hereford bull Li Domino 99375; CHORI-240 Bovine BAC

```

FEATURES

source

Location/Qualifiers

1..846

/organism="Bos taurus"

/mol_type="genomic DNA"

/strain="breed: Hereford"

/db_xref="taxon:9913"

/clone="CH240_345P10"

/sex="Male"

/cell_type="Blood"

/clone_lib="CHORI-240"

/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;

Hereford bull Li Domino 99375; CHORI-240 Bovine BAC

Library (Male) produced by Pieter de Jonga"

Query Match	21.6%	Score 770.7	DB 1	Length 846
Best Local Similarity	95.8%	Pred. No. 0.2		
Matches 801	Conservative 0	Mismatches 34	Indels 1	Gaps 1
355	ATGAACAGTATG-AAAAGGCCAAAATGATAGGATACCTGAAAGAGGAACTCCCCAGGTCACT	413		
846	ATGAACAGTATGAAAAGGCCAAAATGATAGGATACCTGAAAGAGAACTCCCCAGGTCACT	787		
414	AGGTGCCCATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGNAAGAAATGAAGAGAT	473		
786	AGGTGCCCAATGTGCTACTGGAGATCAGTGGAGAAATAACTCCAGNAAGAAATGAAGAGAT	727		
474	GGAGCCAAAGCAAAAAGAAATACCCAGCTGTGATGTGACTGGTGATATAAGCAAGTCCG	533		
726	GGAGCCAAAGCAAAAACAATAACCAGCTGTGATGTGACTGGTGATAGAAGCAAGTCTG	667		
534	ATGCTGTAAAGAGCAATATTGCATAGGAACCTGGAAATGTCAAGTCCATGAATCAAGGCAA	593		
666	ATGCTGTAAAGAGCAGTATTGCATAGGAACCTGGAAATGTCAAGTCCATGAATCAAGGCAA	607		
594	ATTGGAAGTGGTCAACACAGAGATGCCAGAGTGAATGTCAACATCTAGGAATCAGCGA	653		
606	ATTGGAAGTGGTCAACCAAGTGAATGCCAGAGTGAATGTCAACATCTTAGGAATCAGTGA	547		
654	ACTAAAAATGGCATGGAAATGGGTGAAATTTAACTCAGATGACCAATTATATCTACTCTGCGG	713		
546	ACTCAATGGCATGGAAATGGGTGAAATTTAACTCAGATGACCAATTATATCTACTCTGCGG	487		
714	GCAGGAATCCCTCAGAAGAAATGGAGTAGCCATCATGCTCAACAAAAGATCGGAATGC	773		
486	GCAGGAATCCCTCAGAAGAAATGGAGTAGCCATCATGCTCAACAAAAGATCGGAATGC	427		
774	AGTACTTTGGATGCAGTCTCAAAAACACAGAATGATCTCTGTGTGTGTTTCCAAAGGCCAAACC	833		
426	AGTACTTTGGATGCATCTCAAAAATGCAGAAATGATCTCTTTCAITTTCCAAAGGCCAAACC	367		
834	ATTCAATATCACAGTAAATCAAAGTCTATGCCCCAAACCAAGTAATGCTGAAGAACTGAAGT	893		
366	ATTCAATATCACAGTAAATCAAAGTCTATGCCCCAAACCAAGTAATGCTGAAGAACTGAAGT	307		
894	TGNAACGGTCCCTATGAAGACCTTCAAGACCTTTTAGBACTAAACCCCAAAAAGATGTCCT	953		
306	TGAACGGTTCCTATGAAGACCTTCAAGACCTTTTAGBACTAAACCCCAAAAAGATGTCCT	247		
954	TCTCATATTAGGGGACTGGAAATCAAAGTAGGAAGCAAAAGAACCACTCGAGTAAACAGG	1013		
246	TTTCAITATTAGGGCAGCTGGAAATCAAAGTAGGAAGTCAAGAAACCACTGGAGTAAACAGG	187		
1014	CAAAATTGGCCTTTGGAAATACGGAAATGAAGCAGGCAAGAGACTAATAGAGTTTGGCCAGA	1073		
186	CAAAATTGGCCTTTGGAAATACAGAAATGAAGCAGGCAAGAGACTAATAGAGTTTGGCCAGA	127		
1074	AAATGCACTGGTCAATAGCAAAACACCCCTCTTCCAAACAACAAGAGAAGACTCTACACATG	1133		
126	AAATGCACTGGTCAATACAAACACCCCTCTTCCAAATAAAGAGAGAAAACCTCTATACATG	67		
1134	GACATCACAGATGGTCAACACCGAAATACAGATTGATTTATTTCTTTGGACGCCAAA	1189		
66	GACATCACAGATGGTCAACACCGAAATACAGATTGATTTATTTCTTTGGACGCCAAA	11		

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
Bovidae; Bovinae; Bos.
1 (Bases 1 to 947)
Holt, R., Stott, J., Yang, G., Barber, S., Smailus, D., Prabhu, A.-L.,
Tsai, M., Cloutier, A., Lee, D., Ginn, N., Olson, T., Mayo, M.,
Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
Mathewson, C., Wye, N., Vasson, A., Brown-John, M., Jones, S.,
Schein, J., Marra, M., de Jong, P., McWilliam, S., Barris, W.,
Dairymple, B. P. and Tellam, R.
Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
Unpublished (2003)
Other GSSs: CH240_392111.TARBAC13P2
Contact: Rob Holt
Sequencing
The British Columbia Cancer Agency Genome Science Centre
600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
Tel: 604-877-6085
Fax: 604-877-6276
Email: rholt@bcgsc.ca
Clones are derived from the bovine BAC library CHORI-240
(<http://www.chori.org/bacpac/bovine240.htm>). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(<http://www.chori.org/bacpac/ordering.information.htm>). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
British Columbia Genome Sciences Centre, Canada.

FEATURES source

	Query Match	21.5%;	Score 769.501;	DB 1;	Length 947;	
	Best Local Similarity	91.2%;	Pred. No. 0.18;			
	Matches 869;	Conservative	0;	Mismatches 75;	Indels 9;	Gaps 5;
Qy	1829	GCACAGTTAGAACTGGACATGGAAACAACAGACTGGTTCCAACTAGAGAAAGGATATGT	1888			
Db	947	GCACAGTTAGAACTGGACATGGAAATACACACGGGTTCCAAATAGAAAAAGGATATGT	888			
Qy	1889	CAAGGCTGTATATTGTCAACCGGCTTGCTTTAACTCTCATGCAGAG-ACATCATCAGAAAC	1947			
Db	887	CAAGGCTGTATA---TCACCTGTATTATTAACTTATATGCAGAGTACATCATCAGAAAC	831			
Qy	1948	GCTGGGCTGGAAGAAGCACAACTGGAAATCAAGATTCCGGGAGAAATAGCAATAACCTC	2007			
Db	830	GCTGGA CTGGAAGAAGCACAACTGGAAATCAAATTTCCAGGAGAAATATCAGTAACCTC	771			
Qy	2008	AGATATCGAGATGATACCACTTTATGGCAGAAAGTCAAGAGGAACATAAAAGCCTCTTG	2067			
Db	770	AGATATTCAGATGACACACCTTTATGGCAGAAAGTGAAGAGGAAC-TCAAAGCCCCCTTG	712			
Qy	2068	ATGAAGGTGAAAGAGGAGAGTGAAAAAGTTGGCTTTAAAGCTTCAACATTCAAAAAACGAAG	2127			
Db	711	ATGAAGTGAAGTGGAGAGTGAAAAAGTTGGCTTTAAAGCTCAACATTCAAAAAACGAAG	652			
Qy	2128	ATCATGGCATCTGCTCCCATCTCACTTCATPGGAAATAGATGGGAAAACAGTGGAAACAGTG	2187			
Db	651	ATCATGGCATCCGGTATCACACCACTTCATGGCAAAATAGATGGGAAAACAGTGGAAACAGTG	592			

2188 TCAGACTTTATTTTGGGGGCTCAAAATCACTGCAGATGGTGAATGCTGCGCCATGAAA 2247
 Db TCAGACTTTATTTT - GGGGGCTCAAAATCACTGCAGATGGTGAATGCTGCGCCATGAAA 534
 2248 TTAAGAAGACATCTCTCTGGAAGAAAGTTA--ACCACTAGATGACATATTGAAAA 2305
 Db TTAAGAAGATGCTTACTCTCTGGAAGAAAGTTATGACCACTAGACACGCTATTCAAAA 474
 2306 GCAGAGACATTAACCTTGCCAAAGCCCATCTAGTCAAGGCTATGTTTTTCCAGTGG 2365
 Db GCAGAGACATTAACCTTGCCAAAGGCTCATGTAGTCAAGGCTATGTTTTTCCAGTGG 414
 2366 TCATGTATGATGTGAGAGTTGAGCTGTGAAGAAAGCTGAGACCTGAAGAAATGATGCTT 2425
 Db TCATATATGATGTGAGAGTTGAGCTGTGCAAGAGGCTGAGTACCAAGAAATGATGCTT 354
 2426 TTGAAGTCTGGTGTGGAGAGACCTTTGAGAGTCCCTTGACCTGCAAGGAGATCCCAAC 2485
 Db TTGAAGTCTGGTGTGGAGAGACCTTTGAGAGTCCCTTGACCTGCAAGGAGATCCCAAC 294
 2486 AGTCATCTCTGAAGAGATCAGCCCTGGGATTTCTTTGGAAGGAATGATGCTAAAGCTGA 2545
 Db AGTCATCTCTGAAGAGATCAGCCCTGGGATTTCTTTGGAAGGAATGATGCTAAAGCTGA 234
 2546 AACTCCAGTACTTTGGCCACCTGATCAGAGAGCTGACTCTCTGGAAGACCTGATGC 2605
 Db AACTCCAGTACTTTGGCCACCTGATCAGAGAGCTGACTCTCTGGAAGACCTGATGC 174
 2606 TGGAGGGATTTGGGGCAGGAGGAGAGGGGACGACAGAGATGAGATGGCTGGATGGCA 2665
 Db TGGAGGGATTTGGGGCAGGAGGAGAGGGGACGACAGAGATGAGATGGCTGGATGGCA 114
 2666 TCATCACTCTGATGACCTGCTGGTGAACCTCTGGAGTTGGTGGATGACAGAGGAGG 2725
 Db TCATGGCTCAATGACCTGATGCTGAATGAACCTCCAGGATTTGGTGGATGACAGAGGAGG 54
 2726 CCTGTCTCTGGCGGATTCATGGGCTCAAAAGAGTTGGACACGACTCAGCAAC 2778
 Db CCTGGCATGCTGCAATTTATGGGCTCAAAAGAGTCTAGACAGATGAGCGAC 1

RESULT 38
 CC586052 858 bp DNA linear GSS 18-JUN-2003
 DEFINITION CH240_383G13.TARBA13P2 CHORI-240 Bos taurus genomic clone
 CH240_383G13, genomic survey sequence.
 CC586052
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM Bos taurus (cow)
 CC586052.1 GI:31938331
 GSS.
 Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoides;
 Bovidae; Bovinae; Bos.
 1 (Bases 1 to 858)
 Holst, R., Stott, J., Yang, G., Barber, S., Smaluis, D., Prabhu, A., L.,
 Teal, M., Cloutier, A., Lee, D., Gern, N., Olson, T., Mayo, M.,
 Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
 Mathewson, C., Wye, N., Mason, A., Brown-John, M., Jones, S.,
 Schein, J., Warra, M., de Jong, P., McWilliam, S., Barris, W.,
 Dalrymple, B.P. and Tellam, R.
 Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
 Unpublished (2003)
 Other GSSs: CH240_383G13.T7
 Contact: Rob Holt
 Sequencing
 The British Columbia Cancer Agency Genome Science Centre
 600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
 Tel: 604-877-6085
 Fax: 604-877-6276
 Email: rholt@bcgsc.ca
 Clones are derived from the bovine BAC library CHORI-240
 (http://www.chori.org/bacpac/bovine240.htm). For BAC library

availability, please contact Pieter de Jong (pdejong@mail.cho.org).
 Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/ordering/information.htm). This work
 was undertaken as part of the International Bovine BAC Mapping
 Consortium (IBMC) by CSIRO Livestock Industries, Australia and the
 British Columbia Genome Sciences Centre, Canada.
 Plate: 383 row: G column: 13
 Seq primer: SP6
 Class: BAC ends

FEATURES

Location/Qualifiers
 1..858
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /strain="breed: Hereford"
 /db_xref="taxon:9913"
 /clone="CH240_383G13"
 /sex="Male"
 /cell_type="Blood"
 /clone_lib="CHORI-240"
 /note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
 Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
 library (Male) produced by Pieter de Jong"

Query Match 21.5%; Score 769.8; DB 1; Length 858;
 Best Local Similarity 94.9%; Pred. No. 0.2;
 Matches 815; Conservative 0; Mismatches 42; Indels 2; Gaps 2;

QY 289 TGACGAGATGTGGTCCACTGGAGAGGGAAT--GCAAAACCACTTCAGTATTTCTGCTTGA 347
 Db 1 TGACGAGAGTGGTCCACTGGAGAGGGAATGGCAACCACTTCAGTATTTCTGCTTGA 60
 QY 348 GAACCCCATGAACAGTATGAAAGCAAAATGATAGGATCTGAAAGAGGAACCTCCCAAG 407
 Db 61 GAACCCCATGAATAGTATGAAAGCAAAATGATAGGATCTGAAAGAGGAACCTCACCAG 120
 QY 408 GTCACTAGTGTCCCATATGCTACTGGAGATCAGTGGAGAAATACTCCAGAAAGATGA 467
 Db 121 GTTAGTAGTGCCCTATATGCTACTGGAGATCAGTGGAGAAATACTCCAGAAAGATGA 180
 QY 468 AGAGATGGAGCCAAAGCAAAAGAAATACCCAGCTGTGGATGTGATGTGATATAAGCAA 527
 Db 181 AGGATGGAGCCAAAGCAAAAGAAATACCCAGCTGTGGATGTGATGTGATATAAGCAA 240
 QY 528 GGTCCGATGCTGTAAAGACCAATATTCATAGGACCTGGATGTGATGTGATATAAGCAA 587
 Db 241 GGTCTGATGCTGTAAAGACCAATATTCATAGGACCTGGATGTGATGTGATATAAGCAA 300
 QY 588 AGGCCAAATTGGAAGTGGTCAAAACAGAGATGGCAAGAGTGAATGTCAACATTTCTAGGAAT 647
 Db 301 AGGAAATTTGGAAGTGGTCAAAACAGAGATGGCAAGAGTGAATGTCAACATTTCTAGGAAT 360
 QY 648 CAGCGAATTAATGGACTGGAAATGGGTGAATTTAACTCAGATGACCAATTTATCTACTA 707
 Db 361 CAGCGAATTAATGGACTGGAAATGGGTGAATTTAACTCAGATGACCAATTTATCTACTA 420
 QY 708 CTGCGGAGAGAAATCCCTCAGAGAAATGGAGTAGCCATCATGTCAACAAAGAGTCCG 767
 Db 421 CTGCGGAGAGAAATCCCTCAGAGAAATGGAGTAGCCATCATGTCAACAAAGAGTCTG 480
 QY 768 AAATGCATGATTTGGATGACGTCTCAAAAACGACAGAAATGATCTCTGTTGTTCCAAAG 827
 Db 481 AAATGCATGATTTGGATGACGTCTCAAAAACGACAGAAATGATCTCTGTTGTTCCAAAG 540
 QY 828 CAAACCATTCATATCAAGTAATCCAGTCTATGCCCCCAGCAAGTATGCTGAGAGAGC 887
 Db 541 CAAACCATTCATATCAAGTAATCCAGTCTATGCCCCCAGCAAGTATGCTGAGAGAGC 600
 QY 888 TGAAGTTGAACGGTCTTATGAAGACCTTACAGACCTTTTGAACCTTAACACCCAAAAAGA 947
 Db 601 TGAAGTTGAACGGTCTTATGAAGACCTTACAGACCTTTTGAACCTTAACACCCAAAAAGA 660
 QY 948 TGTCTCTCTATTATAGGGACTGGAATGCAAAAGTAGGAGCAAGAAACACCTGAGT 1007

Db 661 TGTCTTTTCATTTATAGGGGACTGGAATGCTAAAGTAGGAAGTCAAGAAACTCTCTGGAGT 720
 QY 1008 AACACGCAAAATTTGGCTTTGGAATACGGAATGAAGACGCGCAAGACTTAATAGAGTTTG 1067
 Db 721 AACAGCAAAATTTGGCTTTGGAATACGGAATGAATAGGCGCAAGACTTAATACAGTTTG 780
 QY 1068 CCAAGAAATGACCTGGTGCATAGCAAAACACCTCTTCCAAACAACACACAGAGAACTCTTA 1127
 Db 781 -CAAGAAATACACTGATCATTAACAACAACACCTCTTCCAAACAACACAGAGAACTCTTA 839
 QY 1128 CACATGGACATCACCAGAT 1146
 Db 840 CACATGGACATCACCAGAT 858

RESULT 39

CC525458/c
 LOCUS CH240_399H6.T7 CHORI-240 Bos taurus DNA linear GSS 17-JUN-2003
 DEFINITION genomic survey sequence.

ACCESSION CC525458

VERSION CC525458.1 GI:31843746

KEYWORDS GSS

SOURCE Bos taurus (cow)

ORGANISM

Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.

REFERENCE

1 (bases 1 to 877)
 Holt, R., Scott, J., Yang, G., Barber, S., Smalhus, D., Prabhu, A.-L.,
 Tsai, M., Cloutier, A., Lee, D., Ginn, N., Olson, T., Mayo, M., Chiu, R.,
 Butterfield, J., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
 Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
 Schein, J., Marra, M., de Jong, P., Keefe, J.W. and Kappes, S.M.
 Bovine BAC End Sequences from Library CHORI-240, PLATES 399 to 478
 Unpublished (2003)

TITLE

Other GSSs: CH240_399H6.TABAC13P2

JOURNAL

COMMENT

Contact: Rob Holt
 Sequencing
 The British Columbia Cancer Agency Genome Science Centre
 600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4B6
 Tel: 604-877-6085
 Fax: 604-877-6276
 Email: rholt@bcgsc.ca

Clones are derived from the bovine BAC library CHORI-240
 (<http://www.chori.org/bacpac/bovine240.htm>). For BAC library
 availability, please contact Pieter de Jong (pdejong@gmail.com).
 Clones may be purchased from BACFAC Resources
 (http://www.chori.org/bacpac/ordering_information.htm). This work
 was undertaken as part of the International Bovine BAC Mapping
 Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
 British Columbia Genome Sciences Centre, Canada.

Plate: 399 row: H column: 6

Seq primer: T7

Class: BAC ends.

FEATURES

Location/Qualifiers

1..877
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /strain="Breed: Hereford"
 /db_xref="taxon:9913"
 /clone="CH240_399H6"
 /sex="Male"
 /cell_type="Blood"
 /clone_lib="CHORI-240"
 /note="Vector: pTABAC1.3; Site 1: MboI; Site 2: MboI;
 Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
 library (Male) produced by Pieter de Jong"

Query Match

Best Local Similarity 21.5%; Score 768.5; DB 1; Length 877;

Matches 839; Conservative 0; Mismatches 46; Indels 7; Gaps 3;

RESULT 40

CC548709/c

LOCUS CH240_432B6.T7

DEFINITION

genomic survey sequence.

ACCESSION CC548709

VERSION CC548709.1

KEYWORDS GSS

SOURCE Bos taurus (cow)

ORGANISM

Bos taurus

Eukaryota; Metazoa;

Chordata; Craniata;

Vertebrata; Euteleostomi;

Mammalia; Eutheria;

Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovidae; Bovinae; Bos.

Bos taurus

Bos taurus

Bos taurus

Bos taurus

344 TTGAGAACCCCATGAACAGTATGAAAAGGC-AAAATGATAGGATCTGAAAAGAGAACTC 402
 Db 877 TTGAGAACCCCATGAACAGTATGAAAAGGC-AAAATGATAGGATCTGAAAAGAGAACTC 818
 QY 403 CCAGGTGAGTGGTGGCCCATATGCTACTGAGATCAGTGGGAAATACCTCCAGAAAG 462
 Db 817 CCAGGTGAGTGGTGGCCCATATGCTACTGAGATCAGTGGGAAATACCTCCAGAAAG 758
 QY 463 AATGAAGATGAGGAGCCCAAGCAAAAGAAATACCCAGCTGTGGATGTGACTGGTATATA 522
 Db 757 AATGAAGGATGAGGAGCCCAAGCAAAAGCAATATCCAGCTGTGGATGTGACTGGTATATA 698
 QY 523 AGCAAGGTCGATGCTGTAAAGAGCAATATTGATAGGAACTCGAATCTGGAATGTGAGTCCATG 582
 Db 697 AGCAAGGTCGATGCTGTAAAGAGCAATATTGATAGGAACTCGAATCTGGAATGTGAGTCCATG 638
 QY 583 AATCAAGGCAAAATTCGAATGGTGTCAACAAGAGATGGCAAGAGTGAATGTCAACATTTCTA 642
 Db 637 AATCAAGGCAAAATTCGAATGGTGTCAACAAGAGATGGCAAGAGTGAATGTCAACATTTCTA 578
 QY 643 GGAATCAGCGAACTAAATGGACTGGATGGTGAATTTTAACTTAACTCAGATGACCATATATC 702
 Db 577 GGAATCAGCGAACTAAATGGACTGGATGGTGAATTTTAACTTAACTCAGATGACCATATATC 518
 QY 703 TACTACTGGGSCAGGAATCCCTCAGAAAGAAATGGAGTAGCCATCATGGTCAACAAAGA 762
 Db 517 TACTACTGGGSCAGGAATCCCTCAGAAAGAAATGGAGTAGCCATCATGGTCAACAAAGA 458
 QY 763 GTCCCAATGAGTACTTGGATGGATGCTCAAAAACGACAGAAATGATCTCTGTTGTTTC 822
 Db 457 GTCTGAATGAGTACTTGGATGGATGCAATCTGAAAACGACAGAAATGATCTCTGTTGTTTC 398
 QY 823 CAAGCAAAACCATTCATATCACAGTAAATCCAAAGTCTATGCCCAACACAGTAATGCTGAA 882
 Db 397 CAAGCAAAACCATTCATATCACAGTAAATCCAAAGTCTATGCCCAACACAGTAATGCTGAA 338
 QY 883 GAAGTGAAGTGAACGCTCTATGAAGACCTTACAGACCTTTTAGAATCAACACCCAAA 942
 Db 337 GAAGTGAAGTGAATGGTCTCTGAAGACCTTACAGACCTTTTAGAATCAACACCCAAA 278
 QY 943 AAAGATGTCCTTCTCATTTATAGGGGACTGGAATCAAAAGTAGGAAGCAAGAAACACCT 1002
 Db 277 AAAGATGTCCTTCTCATTTATAGGGGACTGGAATCAAAAGTAGGAAGCAAGAAACACCT 218
 QY 1003 GGAGTAAACAGCAAAATTTGGCTTGGATAGCGAATGAGGAGGAGGCAAGCAAGCAAGTAAAG 1062
 Db 217 GGAGTAAACAGCAAAATTTGGCTTGGATAGCGAATGAGGAGGAGGCAAGCAAGTAAAG 158
 QY 1063 TTTTCCCAAGAAATGCACTGGTCTATAGCAAAACACCTCTTCCCAACCAACAGAGAGA 1122
 Db 157 TTTTCCCAAGAAATGCACTGGTCTATAGCAAAACACCTCTTCCCAACCAACAGAGAGA 103
 QY 1123 CTCTACATGAGCACTACC-AGATGTCACACCGAAATCAGATTGATTTATTTG 1181
 Db 102 TTCTACCTGGATATCACCAGATGGTCAACCGAAATCAGATTGATTTATTTG 43
 QY 1182 CAGCCAAAGATGGAGAGCTCTATACAGTCAAGCAAAACCAAG 1223
 Db 42 CAGCCAAAGATGGAGAGCTCTATACAGTCTTACAAACCAAG 1

CC548709 841 bp DNA linear GSS 17-JUN-2003
 CH240_432B6.T7 CHORI-240 Bos taurus genomic clone CH240_432B6,
 genomic survey sequence.

ACCESSION CC548709

VERSION CC548709.1

KEYWORDS GSS

SOURCE Bos taurus (cow)

ORGANISM

Bos taurus

Eukaryota; Metazoa;

Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovidae; Bovinae; Bos.

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus


```

FEATURES
  source
    Class: BAC ends.
    Location/Qualifiers
      1..865
        /organism="Bos taurus"
        /mol_type="genomic DNA"
        /strain="breed: Hereford"
        /db_xref="taxon:9913"
        /clone="CH240_412K21"
        /sex="Male"
        /cell_type="Blood"
        /clone_lib="CHORI-240"
        /note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
        Hereford bull Li Domino 99375; CHORI-240 Bovine BAC
        library (Male) produced by Pieter de Jong"

Query Match      21.5%; Score 768.4; DB 1; Length 865;
Best Local Similarity 94.5%; Pred. No. 0.2;
Matches 817; Conservative 0; Mismatches 46; Indels 2; Gaps 2;

QY 202 CACTAGGACACAGCCTTGTCTTAATCAATCAATGAAA-CTAAGCCATGCCCGTGGGGCAACCC 260
DB 865 CACTAGCACCACAGCCTTGTCTTAATCAATCAATGAAAACCTAAGCCATGCCCGTGGGGCAACCC 806
QY 261 AGATGGGAGGTCTATGTGGAGAGATCTGACAGATGTGTCTCACTGGAGAGGGAATG 320
DB 805 AAGACAGGGGGTCTATGTGGAGAAATCTGACAGAAATGTGTCTCACTGGAAAGGGAATG 746
QY 321 -CAAAACCACTTCAGTATTTCTTGCCTTTGAGAACCCCATCAACAGTATGAAAAGGCCAAATG 379
DB 745 TCBAACCACTTCAGTATTTCTTGCCTTTGAGAACCCCATGACAGTATGAAAAGGTTAATG 686
QY 380 ATAGATACTGAAGAGGNACTCCCGAGTCCAGTACGTAGTGGCCCATATGCTACTGGAGATC 439
DB 685 ATAGATACTGAAGAGGNACTCCCGAGTCCAGTACGTAGTGGCCCATATGCTACTGGAGATC 626
QY 440 AGTGAGAAATTAACCTCCAGAAAGATGAAGAGATGGAGCCAAAGCAAAAGAAATACCCAG 499
DB 625 AGTGAGAAATTAACCTCCAGAAAGATGAAGAGATGGAGCCAAAGCAAAAGAAATACCCAG 566
QY 500 CTGTGATGTGACTGGTGATTAAGCAAGGTCCGATGCTGTGTAAGAGCAATATTGATAG 559
DB 565 CTGTGATGTGACTGGTGATTAAGCAAGGTCCGATGCTGTGTAAGAGCAATATTGATAG 506
QY 560 GAACCTGGAATGTCAAGTCCATGAATCAAGGCAAAATGCAAGTGGTCAACCAAGAGATGG 619
DB 505 GAACCTGGAATGTCAAGTCCATGAATCAAGGCAAAATGCAAGTGGTCAACCAAGAGATGG 446
QY 620 CAAGATGAATGTCAACATTTAGGAATCAGCGAATCAAAATGGAATGGAAATGGTGAAT 679
DB 445 CAAGATGAACACTGACATTTCTAGAAATCAGCAAACTCAAGTGGACTGGAATGGTGAAT 366
QY 680 TTAACCTCAGATGACATTTATATCTACTCTGGGGCAGCAATCCCTCAGAGAAATGGAG 739
DB 385 TTAACCTCAGATGACATTTATATCTACTCTGGGGTGAATCCCTCAGAGAAATGGAG 326
QY 740 TAGCCATCATGTCAACAAAGAGTCCGAAATGCAATGCTGGATGCGAGTCTCAAAACG 799
DB 325 TAGCCATCATGTCAACAAAGAGTCCGAAATGCAATGCTGGATGCGAATCTCAAAACG 266
QY 800 ACAGATGATCTCTGTTTGTTCCTCAAGGCAACCAATCAATATCACAGTATCAAGTCT 859
DB 265 ACAGATGATCTCTGTTTGTTCCTCAAGGCAACCAATCAATATCACAGTATCAAGTCT 206
QY 860 ATGCCCCAACAGTAAATGCTGAAGAGCTGAAGTTGAACGGTCTCTATGAAGACCTCAAG 919
DB 205 ATGCCCCAACAGTAAACGCTGAAGAGCTGAAGTTGAACAGTCTCTATGAAGACCTCAAG 146
QY 920 ACCTTTTAGAATCAACACCAAAAGATGCTCTTCTCATTTATAGGGGACTGGGAATGCAA 979
DB 145 ACCTTTTAGAATCAACACCAAAAGATGCTCTTCTCATTTATAGGGGACTGGGAATGCAA 86
QY 980 AAGTAGGAAGCAAGAAACACCTGAGTAACAGGCAAAATTTGGCTTGGATACGGAATG 1039

```

```

DB 85 AAGTAGGAAGTCAAGAAACACCTGGAGTAACAGGCAAAATTTGGCCTTAGAATACAGAATG 26
QY 1040 AAGCAGGGCAAGACTAATAGATT 1064
DB 25 AAGCAGGGCAAGACTAATAGATT 1

RESULT 42
CC581625/c
LOCUS
DEFINITION
  CH240_376F14.T7 CHORI-240 Bos taurus genomic clone CH240_376F14,
  genomic survey sequence.
ACCESSION
  CC581625
VERSION
  CC581625.1 GI:31929477
KEYWORDS
  GSS.
SOURCE
  Bos taurus (cow)
  ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
    Bovidae; Bovinae; Bos.
  REFERENCE
    1 (bases 1 to 822)
    Holt, R., Stott, J., Yang, G., Barber, S., Smalley, D., Prabhu, A.-L.,
    Tsai, M., Cloutier, A., Lee, D., Ginn, N., Olson, T., Mayo, M., Chiu, R.,
    Butterfield, J., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
    Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
    Schein, J., Marra, M., de Jong, P., McWilliam, S., Barris, W.,
    Dalrymple, B.P. and Tellam, R.
    Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
    Unpublished (2003)
    Other GSSs: CH240_376F14.TARBAC13P2
    Contact: Rob Holt
  Sequencing
    The British Columbia Cancer Agency Genome Science Centre
    600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4B6
    Tel: 604-877-6085
    Fax: 604-877-6276
    Email: rholt@bccsc.ca
    Clones are derived from the bovine BAC library CHORI-240
    (http://www.chori.org/bacpac/bovine240.html). For BAC library
    availability, please contact Pieter de Jong (pdejong@email.cho.org).
    Clones may be purchased from BACPAC Resources
    (http://www.chori.org/bacpac/ordering_information.htm). This work
    was undertaken as part of the International Bovine BAC Mapping
    Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
    British Columbia Genome Sciences Centre, Canada.
    Plate: 376 Row: F Column: 14
    Seq primer: T7
    Class: BAC ends.
    Location/Qualifiers
      1..822
        /organism="Bos taurus"
        /mol_type="genomic DNA"
        /strain="breed: Hereford"
        /db_xref="taxon:9913"
        /clone="CH240_376F14"
        /sex="Male"
        /cell_type="Blood"
        /clone_lib="CHORI-240"
        /note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
        Hereford bull Li Domino 99375; CHORI-240 Bovine BAC
        library (Male) produced by Pieter de Jong"

Query Match      21.5%; Score 767.3; DB 1; Length 822;
Best Local Similarity 96.6%; Pred. No. 0.21;
Matches 794; Conservative 0; Mismatches 27; Indels 1; Gaps 1;

QY 408 GTCAGTAGTGCCCATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGATGA 467
DB 822 GTCAGTAGTGCCCATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGATGA 763
QY 468 AGAGATGGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 527
DB 762 AGGATGGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 703

```



```

Db      540  AGTAGGAGTCAAGAACACCTGGGGTAACAGGCAAAATTTGGGCTTGGATATGGAAATG 599
QY      1040  AGCAGGGCAAGACTAATAGAGTTTTCAGAGAAATGCACTGGTCATAGCAACACCC 1099
Db      600  AAGCAGGGCAAGGCTAATAGAGTTTTCAGAGAAATGCACTGGTCATAGCAACACCC 659
QY      1100  TCTTCCAAACACACAGAGAGACTCTACATGAGATCAGATGAGATGGTCACACCGAA 1159
Db      660  TCTTCCAAACACACAGAGAGACTCTACATGAGATCAGATGAGATGGTCACACCGAA 719
QY      1160  ATCAGATTGATTATATCTTTTCAGCCAAAGATGGAGAGCTCTATACAGTCAGCAAAAA 1219
Db      720  ATCAGATTGATTATATCTTTTCAGCCAAAGATGGAGAGCTCTATACAGTCAGCAAAAA 779
QY      1220  CAAGACC-AGGAGCTTACTGTGGCTCAGATCATGAACTCTTATTCGCAATTCAGACTT 1278
Db      780  CAAGCCGTGGAGCTGCTGTGGCTCAGATCATGAACTCTTATTCGCAATTCAGACTT 839
QY      1279  AAATTGAAGAAAGTAGGAAAA 1300
Db      840  AAATTGAAGAAAGTAGGAAAA 861

```

RESULT 44

CC512505

LOCUS

CH240_356L21.TARBAC13P2 CHORI-240 Bos taurus linear GSS 17-JUN-2003
 CH240_356L21, genomic survey sequence.

ACCESSION

CC512505

VERSION

GSS:31830793

KEYWORDS

GSS

SOURCE

Bos taurus (cow)

ORGANISM

Bos taurus

REFERENCE

1. (bases 1 to 870)
 Holt, R., Stott, J., Yang, G., Barber, S., Smallos, D., Prabhu, A.-L.,
 Tsai, M., Cloutier, A., Lee, D., Ginn, N., Olson, R., Mayo, M.,
 Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
 Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
 Schein, J., Marra, M., de Jong, P., McWilliam, S., Barris, W.,
 Dalrymple, B.P., and Tellam, R.
 Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
 Unpublished (2003)
 Other_GSSs: CH240_356L21.T7
 Contact: Rob Holt

TITLE

The British Columbia Cancer Agency Genome Science Centre

JOURNAL

600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4B6

COMMENT

Sequencing
 The British Columbia Cancer Agency Genome Science Centre
 600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4B6
 Tel: 604-877-6085
 Fax: 604-877-6276
 Email: rholt@cgsc.ca
 Clones are derived from the bovine BAC library CHORI-240
 (http://www.chori.org/bacpac/bovine240.htm). For BAC library
 availability, please contact Pieter de Jong (pdejong@mail.cho.org).
 Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/ordering/information.htm). This work
 was undertaken as part of the International Bovine BAC Mapping
 Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
 British Columbia Genome Sciences Centre, Canada.
 Plate: 356 row: L column: 21
 Seq primer: SP6
 Class: BAC ends.

FEATURES

Location/Qualifiers

```

1..870
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="bred: Hereford"
/db_xref="taxon:9913"
/clone="CH240_356L21"
/sex="Male"

```

```

Query Match      21.5%; Score 766.4; DB 1; Length 870;
Best Local Similarity 93.9%; Pred. No. 0.2;
Matches 818; Conservative 0; Mismatches 51; Indels 2; Gaps 2;

QY      289  TGACAGAATGTGGTCCACTGGAGAGGGAAT-GCAAAACCACTTCAGTATTCTTGGCCTTGA 347
Db      1    TGACAGAAGGTGGTCCGCTGGAGAGGGAATGGCAATCCACTTCAGTATTCTTGGCCTTGA 60
QY      348  GBAACCCCATGACAGTATGAAAGCGCAAAATGATAGGATCTGAAAGAGAACTCCCCAG 407
Db      61  AAACCCCATGACAGTATGAAAGCGCAAAATGATAGGATCTGAAAGAGAACTCCCCAG 120
QY      408  GTCACTAGGTGCCCCATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGATGA 467
Db      121  GTCACTAGGAGCCCAATATGCTACTGGATATCAGTGGAGAAATAACTCCAGAAAGATGA 180
QY      468  AGAGATGAGCGCAAGCAAAAGAAATACCCAGCTGTGGATGTGCTGGTATATAGCAA 527
Db      181  AGGATGAGCGCAAGCAAAAGAAATACCCAGCTGTGGATGTGCTGGTATATAGCAA 240
QY      528  GGTCCGATGCTGTAAGAGCAATATTGCTAGGAACCTGGAATCTCAGGTCCATGAATCA 587
Db      241  GGTCTGTGCTGTAAGAGCAATATTGCTAGGAACCTGGAATCTCAGGTCCATGAATCA 300
QY      588  AGGCAAAATGGAAGTGGTCAAAAGAGATGGCAAGAGTGAATGTCAAACTTCTAGGAAT 647
Db      301  AGGCAAAATGGAAGTGGTCAAAAGAGATGGCAAGAGTGAATGTCAAACTTCTAGGAAT 360
QY      648  CAGCGAACTAAATAGGACTGGAATGGGTGAATTTAACTCAGATGACCATTTATCTACTA 707
Db      361  CAGCGAACTAAATAGGACTGGAATGGGTGAATTTAACTCAGATGACCATTTATCTACTA 420
QY      708  CTGCGGGCAGAAATCCCTCAGAGAAATGAGTAGCCATCATGTGTAACAAAAGAGTCCG 767
Db      421  CTGCGGGCAGAAATCCCTCAGAGAAATGAGTAGCCATCATGTGTAACAAAAGAGTCCG 480
QY      768  AAATCAGTACTTGGATGCACTCTCAAAAACGACAGAAATGATCTCTGTTTGTTCCAAGG 827
Db      481  AAATCAGTACTTGGATGCACTCTCAAAAACGACAGAAATGATCTCTGTTTGTTCCAAGG 540
QY      828  CAAACCATTCATATCAGTAATCCAAAGTCTATGCCCAACCAAGTAATGCTGAAGAGC 887
Db      541  CAAACCATTCATATCAGTAATCCAAAGTCTATGCCCAACCAAGTAATGCTGAAGAGC 600
QY      888  TCAAGTTGAACGGTCTCTATGAAGACCTTACAAAGACCTTTTAGAAGTAAACCCCAAAAGA 947
Db      601  TGAAGTGAACGGTCTCTATGAAGACCTTACAAAGACCTTTTAGAAGTAAACCCCAAAAGA 659
QY      948  TGTCTCTCTATATAGGGGACTGGAAATGCAAAAGTAGGAAGCAAAAGAACACCTGGAGT 1007
Db      660  TGTCTCTCTATATAGGGGACTGGAAATGCAAAAGTAGGAAGTCAAGAAACACCTGCAGT. 719
QY      1008  AACAGGCAAAATTTGGCTTGAATACGGAATGAAGCAGGGCAAGAGCTAATAGAGTTTG 1067
Db      720  AACAGGCAAAATTTGGCTTGAATACGGAATGAAGCAGGGCAAGAGCTAATAGAGTTTG 779
QY      1068  CCAAGAAATGCACTGGTCTATAGCAAAACACCTCTTCCAAACACCAAGAGAGAGACTCTA 1127
Db      780  CCAAGAAATGCACTGGTCTATAGCAAAAGACCTTTTCCAAACACCAAGAGAGAGACTTTA 839
QY      1128  CACATGGACATCACCAGATGGTCAACACCGA 1158
Db      840  CACATGGACATCACCAGATGGTCAACACCGA 870

```

RESULT 45
 CCS40150

```

LOCUS      CC540150              862 bp      DNA      linear      GSS 17-JUN-2003
DEFINITION CH240_420F7.TARBA13P2 CHORI-240 Bos taurus genomic clone
ACCESSION  CC540150
VERSION    CC540150
KEYWORDS   GSS.
SOURCE     Bos taurus (cow)
ORGANISM   Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
            Bovidae; Bovinae; Bos.
REFERENCE  1 (bases 1 to 862)
AUTHORS   Holt, R., Stott, J., Yang, G., Barber, S., Smalhus, D., Prabhu, A.-L.,
            Tsai, M., Cloutier, A., Lee, D., Gryn, N., Olson, T., Mayo, M.,
            Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
            Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
            Schein, J., Marra, M., de Jong, P., Keele, J. W. and Kappes, S. M.
            Bovine BAC End Sequences from Library CHORI-240, PLATES 399 to 478
            Unpublished (2003)
            Other GSSs: CH240_420F7.T7
            Contact: Rob Holt
            Sequencing
            The British Columbia Cancer Agency Genome Science Centre
            600 W. 10th Ave. Vancouver, British Columbia, Canada V5Z 4E6
            Tel: 604-877-6085
            Fax: 604-877-6276
            Email: rholt@bcgsc.ca
            Clones are derived from the bovine BAC library CHORI-240
            (http://www.chori.org/bacpac/bovine240.htm). For BAC library
            availability, please contact Pieter de Jong (pjejong@mail.cho.org).
            Clones may be purchased from BACPAC Resources
            (http://www.chori.org/bacpac/ordering_information.htm). This work
            was undertaken as part of the International Bovine BAC Mapping
            Consortium (IBMC) by CSIRO Livestock Industries, Australia and the
            British Columbia Genome Sciences Centre, Canada.
            Plate: 420 row: F column: 7
            Seq primer: SP6
            Class: BAC ends.
FEATURES   Location/Qualifiers
            source          1..862
                        /organism="Bos taurus"
                        /mol_type="genomic DNA"
                        /strain="breed: Hereford"
                        /db_xref="taxon:9913"
                        /clone="CH240_420F7"
                        /sex="Male"
                        /cell_type="Blood"
                        /clone_lib="CHORI-240"
                        /notes="vector: PTARBAC1.3; Site 1: MboI; Site 2: MboI;
                        Hereford bull Li Domino 99375; CHORI-240 Bovine BAC
                        library (Male) produced by Pieter de Jong"
Query Match      21.4%; Score 765.9; DB 1; Length 862;
Best Local Similarity 94.3%; Pred. No. 0.2;
Matches 815; Conservative 0; Mismatches 46; Indels 3; Gaps 2;

QY 320 GCAAAACACATTCAGTATTCCTTGCTTGAGAACCCCATGAACAGTATGAAAGGCAAAATG 379
DB 1 GCAATACATTCAGTATTCCTTGCTTGAGAACCCCATGAACAGTATGAAAGGCAAAATG 60
QY 380 ATAGGATCTGAAAGAGGAAGT-CCCCAGGTCTAGTGGTCCCCCATATCTACTGGAGAT 438
DB 61 ATAGGATCTGAAAGAGGAAGTCCCCAGGTCTAGTGGTCCCCCATATCTACTGGAGAT 120
QY 439 CAGTGGAGAAATACCTCCAGAAAGATGAAGAGTGGAGCCCAAGCAAAAGAAATACCCA 498
DB 121 CAGTGGAGTAACTACCTCCAGAAAGATGAAGAGTGGTGGCCCAAGCAAAAGAAATACCCA 180
QY 499 GCTGTGGATGTGCTGCTGATATAGCAAGGTCGATGCTGTAAAGAGCAATATTGCGATA 558
DB 181 GCTGTGGATGTGCTGCTGATATAGCAAGGTCGATGCTGTAAAGAGCAATATTGCGATA 240
QY 559 GGAACCTGGAAATGTCAGGTCCATGATCAAGGCAAAATTCGAAGTGGTCAAACAGAGATG 618

```

```

DB 241 GGAACCTGGAAATGTTAGTCCATGATCAAGGCAAAATTCGAAGTGGTCAAACAGAGATG 300
QY 619 GCAAGAGTGAATGTCAACATTTCTAGGAATCAGCGAACTAAAAATGGCAATGGGTGAA 678
DB 301 GCAAGAGTGAATGTGACATTTCTAGGAATCAGCGAACTAAAAATGGCAATGGGTGAA 360
QY 679 TTAACTCAGATGACCAATTTATCTACTCTGCGGCGAGGAATCCCTCAGAAGAAATGGA 738
DB 361 TTAACTCAGATGACCAATTTATCTACTCTGCGGCGAGGAATCCCTCAGAAGAAATGGA 420
QY 739 GTAGCCATCATGGTCAACAAAAAGAGTCCGAAATGCAGTACTTTGGATGTCAGTCTCAAAAAC 798
DB 421 GTGGCCATCATGGTCAACAAAAAGAGTGTAGAAATGCAGTACTTTGGATGTCAGTCTCAAAAAT 480
QY 799 GACAGNAATGATCTCTGTTGTTTCCAGGCAAAACCAATTCATATACAGTATCCAAAGTC 858
DB 481 GACAGNAATGATCTCTGTTTCAATTTCCAGGCAAAACCAATTCATATACAGTATCCAAAGTC 540
QY 859 TATGCCCCCAACCAAGTAAATGCTGAAGAACTGAAGTTGAACGTCCTATGAAGACCTACAA 918
DB 541 TATGCCCCCAACCAAGTAAATGCTGAAGAACTGAAGTTGAATGTTCTATGAAGAACTACAA 600
QY 919 GACCTTTTAGAACTAACACCCCAAAAAGATGTCCTTTCTCATATAGGGGACTGGAATGCA 978
DB 601 GACCTTTTAGAACTAACACCCCAAAAAG--GTCTCTTTTCATTATAGTGTGACTGGAATGCA 658
QY 979 AAAGTAGGAAGCAAGAAACACCTGGAGTAAACAGGCAAAATTTGGCTTGGATACGGAT 1038
DB 659 AAAGTAGGAAGTCAAGAAACACCTGGAAATAACAGGCAAAATTTGGCTTGGAGTACAGACT 718
QY 1039 GAAGCAGGGCAAGAACTAATAGAGTTTTGCCAAGAAATGCACTGGTTCATAGCAAAACACC 1098
DB 719 GAAGCAGGGCAAGAGCTAATAGAGTTTTGGCCAAGAGAGACACTGGTTCATAGCAAAACACC 778
QY 1099 CTCCTTCCAAACACACAGAGAGAGTCTACACATGGACATCCAGAGATGGTCAACACCGA 1158
DB 779 CTCCTTCCAAACACACAGAGAGAGTCTCTATACATGGACATCCACCTGATGGTCAACACCGA 838
QY 1159 AATCAGATTGATTATATTTCTTTC 1182
DB 839 AAACAGATGGATTATATTTCTTTC 862

```

Search completed: August 24, 2004, 13:41:22
Job time : 79 secs


```
QY 530 AGCTTTGTGTTTG-GTGAATAGTCTGTAATATCT-CTAGTCCACTTGGTTATGA 587
Db 53696 ATTCTTTTGTCTTAGAATTGAATGTTCTATAAATATCTGTAGATCCATTTGGTTTCATAG 53637
QY 588 CATCAGTTAGCTCCAGCAATTTCTGTTTCTGTTTCTGTTTGTGAGATGACCTAACTGTTGGA 647
Db 53636 CATCTGTTACTCTTCACTGCTGCTCTGTTTAG-TTTGTGTTTCCAAGATCTGCTCAATGCT 53578
QY 648 GAGAAATGGGTAATGAAGTAGCCCACTATCTGCTGTGAGGT-CAATATGTGATTTTACG 706
Db 53577 GAGAGTAGGGTGTGGAAGTCTCCCACTATTATTGTGTGCTGCAAGTGTGCTTTCAGC 53518
QY 707 TGTAGCTGTGTTGTTTATGAACCTTGGTGACATTTGTTTGTGATGACATGACATTAAGA 766
Db 53517 TTTAGTAAAGTTTCTTTTATGAATGTTGGTGCCCTTGCATTTGGAGATAGATGTTTCTGAGA 53458
QY 767 ATTGCAATGTCCTCTTGTGTGA-TTTTCTTGTGATGCCCTATGTAGTATTTCTTCCCAATCT 825
Db 53457 ATTTAGAATGCATCTTGTGTAGATTTTCTTGTGATGAGATGAAGTATCTTCTCTTATTC 53398
QY 826 CA---TCTGCTTGTGTTTGGGTTTAAGTCTATTAGTACAGATTTAAATGACTGTAT--- 879
Db 53397 TATTTTGTGAACCTTTGGTTGAAGTGTGATTTTATTGATACCTAGAAATGAATGATCT 53338
QY 880 -CGGCTTCTTCTAGGGCCATTTGCTTAGAATA-TCTTTTCCATCTCTTTTACTCTAAGG 937
Db 53337 CCAGCTTGTCTTCTGGGACCAATTTGCTTGGAAATGTTTCAAGCCCTTTACTCTGAGG 53278
QY 938 -----TGATGTCTATCCATGGTAGGTTGTCTTTTGGATGCGACGAGTAGGATGATCT 991
Db 53277 TAGTACTGTCTCTTGTATCTAGTGGTGCATTTCTGTATGCGCAAAATGCTGGGTCC 53218
QY 992 TGTTTTCATATCCATCTGTTTACCAGATCTTTTCTAGAGAAATTAAGATCATTTAGT 1051
Db 53217 TACTTACATATCCAGTCTGTAGTCCATGCTCTTTTATTAGAGATTCAG-----TC 53166
QY 1052 CATTCATGTTGAGAAATTAATGAGCAGTGTGTGTGGAATCTTTGTATCTTGCACTGT 1111
Db 53165 CTTTCATGTTAAGAGATATTAAGGAAAGTGTATGTTTCTTCTCTATTTTGTGTTA 53106
QY 1112 GAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1171
Db 53105 GAGTGGAAATATGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 53063
QY 1172 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1231
Db 53062 -----T 53062
QY 1232 TCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1291
Db 53061 GAAAGAAGATACCTTTCTGCTTTTCTAGGGTGTAGTTTCCCTCCTTGTGTGAGATTTT 53002
QY 1292 TCTCTAGCTCTCT--TAGGCTCTGATTTGAAGATAGATATCTTTACATCTGATTTTA 1349
Db 53001 TTGCTGTGTTATCTTTGTAGGCTGATTTGTGAAGAATATGTTGAATTTGTTTGT 52942
QY 1350 TCTTGAATGCTCTTTCTTCTCCAACTATTTGTGACAGAAAGTTTCTTAAGTCAGTATG 1409
Db 52941 TCATGGAATATATGTTTCTACTGTCTATGTTAAATTGAGATTTTGTGCGGTATAGTAGC 52882
QY 1410 CTGGCTGACATCTGTAGTCTCTTGGAGTCTGTAGCAGATCTGTGCGAGGCTCTTACA 1469
Db 52881 CTGGAGTATATTTGTGTTCTCTACGGGCTGTATGAGATCTGCTAAGATCTCTTAGCT 52822
QY 1470 TTTTGAAGTTCTAATGGAAAGTCAAGTGTAAATCTTAATACATCTGCTCTTATATGTTAA 1529
Db 52821 TTCACAGTCTCT-TGTGAGAGTCTGTTTAAATCTGAGAGTCTGCGCTTATATGTTAC 52763
QY 1530 TTGGCTCTTTTCCCTTGCATCTTTTAAATCTTTCTTTTGTCTATACCTTTTGTGATTT 1589
Db 52762 TT-GACCTTTTCCCTTACTGCTTTTAAATATATCTTTCTTTGTTTGTGCAATTTGTGTTT 52704
```

```
QY 1590 TGATTTATATGCACTGTGGGAGTTTCTTTTCCGGTCCAATCTATTTTGGTGTGTTTGTATG 1649
Db 52703 TGATTTATATGATGATGGGAGAAATTTCTTTTCTGGTCCAGTCTATTTTGGAGTTCTGTAGC 52644
QY 1650 CTTCTTGTACCTTGTAGGACATCTCTTCTCAAGGTTAGGAAATTTTCTTTT---TTG 1705
Db 52643 CTTCTTGTATGTTTATGGGCACTCTCTTCTTTAGGTTAGGAAATTTTCTTCTATACATTT 52584
QY 1706 GTTTCTTGAATAATTTTCCCTGCTTTTGACCTGCCTTCTTCCCTTCC-----TCT 1758
Db 52583 TTTTGTGGAAGATATTTACTGACCTTTTAAGTTGTAATTTTCTATACCTTAATATCC 52524
QY 1759 ATTCTCTTGGTTTTTGCATAGTCTCTGCTTCTGCTTCTGGAATTTTATGCTCGAATTAATTT 1818
Db 52523 TTAGATTTGGTCTTCCATTTGCTCTGATTTCTGGAATTTTGGAGTTAGGAGCTTTT 52464
QY 1819 TAGACTTACATTTCTTTTGACCAAGGATCATTTCTTCTATCTTCTATCTTCTTCACTGCTG 1878
Db 52463 TGCATGTTGTGTTTTCTTGTACTGTTCTCTTGATGTTTTTATGGAATTTTCTGCACTG 52404
QY 1879 AGATTCCTCTCTCTATCTCTGTTATCTCTCACTGAGGCTTGTCTCTGAGGTTCTCTG-TT 1937
Db 52403 AGATTCCTCTCTCTATCTCTGTTATCTCTGTTATCTCTGCGATGCTTGCATCTATAACTCCTGATC 52344
QY 1938 GGCTTCTTAATTTTTCATTTCCAGATTTCTTCACTTGGGTTTTGTTTATTAATCTTA 1997
Db 52343 TCTTCTAGGTTTTCTAACTCCAGGCTGTCTCTCTTTGTATCTTCTTTATTTGTTTCTTA 52284
QY 1998 TTTCCACTTTCAGTCTCTGAATGTTTTTACTCATTTTCTCT---CCAGTATTTACATTTT 2054
Db 52283 TTTCTCTTTTATGATCTTGGATGTTTTGTTTCAATTTCCCTTGGCTATTTGATTTGTTTTT 52224
QY 2055 CATAGGTTTCTTTAATGGAATTTATTCATTTCTCTTCT 2091
Db 52223 CCTGTCACTTCTTAAGGATTTTGTGTTGTGTTTC 52187
```

RESULT 35

```
US-10-322-281-495
; Sequence 495, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 52945201000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 495
; LENGTH: 176594
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(176594)
; OTHER INFORMATION: n = A,T,C or G
US-10-322-281-495
```

Query Match

```
Best Local Similarity 24.5%; Score 558.6; DB 1; Length 176594;
Matches 1149; Conservative 0; Mismatches 559; Indels 130; Gaps 19;
QY 292 TCTATTTCTTGATTTCTATCTTGGCTCATTTTAACTCAGTAGTGTGTTGGTTTCC 351
Db 129540 TTTCTTTCTTTATTCCTTCTTCCCAAGGATCATTTGAGAAGAGTGTGTTCCGTTTCC 129599
QY 352 AATAGTTTGAAGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 411
Db 129600 ACGTGAAGTTGGCTTTCCCAATCTATCTATGTTGTTATGAAGATC---AGCCTTAGTCCAT 129656
QY 412 GGTGGTCAGATAGACATAGATATTAATTTCAATGTCTTTTATCTGTCGAGACTTGCTT 471
```

Db 129657 GGTGATCTGATAGGATGCGACAAATTTCAATATTTTGTATCTGTGTAGGCGCTGTT 129716
QY 472 TGTGTTTGAATATGTTTCAATTTTGGAGA--GTTTCATAGGCTGCTGACAGAGGTAC 529
Db 129717 TGTGACCAATATATAGTCAATTTTGGAGAGGTACCATGAGGTGCTGAGAGAAATGTAT 129776
QY 530 AGTCCTTTGTTGTTTGG-TGAAATAGTCTCTGTAATATCT-CTAGGTCCCACTTGGTTTATGA 587
Db 129777 ATCCCTTTGGTTTGGATATAAAATGTTCTAGATATCTGTTCAATCTATTTGTTTCAAA 129836
QY 588 CATCAGTGTAGTCCAGCAATTTCTGTTTTCGTTTGTGTTTGTGATGACCTAACTGTTGGA 647
Db 129837 CTTCTGTTTGTAGTTTCACTGTGTCCCTGTTTAG-TTTCTGTTTCCATGATCTGTTCATTGGT 129895
QY 648 GAGATGGGGTATTTGAAGTAGCCCACTATCTGTGTGTAGGT-CAATATGTGATTTTAGC 706
Db 129896 GAGAGTGGGGTGTGAAATCTCCCACTATTTATTTGTGTGATGCAATGTGTGTTGAC 129955
QY 707 TGTAGCTGTCTGTTGTTTGAACCTTGGGTGACATGTTGTTGGTGTGACATAGACATTAAGA 766
Db 129956 TTTACTAAAGTTTCTTTTAAATAATGTGGCTGCCCTTGCAATTTGGAGCATAGATATTCAG 130015
QY 767 ATTGCAATCTCTCTTGGTGGATTTT-CCTTTGATGCTCTATGTAGTATTTCTTCCAAATCT 825
Db 130016 ATTGAGAGTCTCTTGGAGATTTTACCTTTGATGAGTATGAAGTGTCCCTCTCTTGTCT 130075
QY 826 CATCTGCTTAGTTTGGGTTTAAGTCTA-TTAGTCCAGATATTAATAAGTCTGTATCGGCT 884
Db 130076 TTTTGTGATAACTTTTGGTTGGAAGTCGATTTTATTCAATATTAGAATGGCTACTCCACCT 130135
QY 885 TCGTCTTAGGCGCAATTTGCTTAGATA-TCTTTTCCATCTCTTTTACTCTAAGGTATGT 943
Db 130136 TGTGTTCTTCCAGCCCAATTTGCTTGGAAAAATGTTTCCAGCCCTTCAATCTGAGGTAGGT 130195
QY 944 CTATCCAT---GGTAGGTGTCTTTTGGATGACAGCAGTAGGATGGATCTTGTTTTTCAT 1000
Db 130196 CTGTCTTTTCCCTCGGATGGGTTTCTGTAAAGCAGCAAAATGTTGGGTCTGTTGTGT 130255
QY 1001 ATCCATCTGTTTACCAGTATCTTTTCTAGAGAAATTAAGATCAATGAGTCATGTATGT 1060
Db 130256 AGCCAGTCTGTAGTCTATGTCTTTTATTTAGGGAATTTGAG-----TCCATTTGATAT 130307
QY 1061 TCAGAAATATCAATGAGCAGTGTGTTGATTTCTTGTATCTTGCACCTGTGGAAGTGT 1120
Db 130308 TAAGAGATATTAAGAAAGTAATTTGTTGCTCTCTATTTTGTGTTAAAGTTGCA 130367
QY 1121 GT 1180
Db 130368 TTTCTGTTCTGTGGCTGTCTTCTTTTAGGTTTGT----- 130400
QY 1181 GTGTGTGTGTGTGTGTGTGTGTCTCTCCCTCTTTTGATTTTGGCCTGGAATAT 1240
Db 130401 -----TTGAAGGAT 130409
QY 1241 TTAATTTATCATATTTCTTGAATGTGGTAACATCTTTAGATTTGAAGTTTTCCTTAGC 1300
Db 130410 TACTTTCTGTCTTTCTAGGGTGTCTTCCATCTTTTATTTGTTTCTTTTCTTCT 130469
QY 1301 CTCTTT-----TAGTCTGCATTTGAAGATAGATATCTTTACATCTGATTTTATCT 1352
Db 130470 GTTGTATCTCTTTGAAGGGCTGGATTCATCGAAAGATAAATGTGAAATTTGGTTTGTCTG 130529
QY 1353 TAGAATGTCTTTCTTTCTCCCACTATTTGTGACAGAAAGTTTCTTAAGTGTGAGTGTCTG 1412
Db 130530 TGGATATCTTTGGTTTCTTCCATCTATGTATTTGAAGTTTGGCTGGGTATAGTGTCTG 130589
QY 1413 GCCTGACATCTGTAGTCTCTTGGAGTCTGTAGACATCTGTGCAAGGCGCTTCTTACATTT 1472
Db 130590 GGCTGGCAATTTGTGTCTCTTGTAGTGTCTGTATAACATCTGTCCAGGCTCTTCTGCGCTTC 130649
QY 1473 TGAGTTCTTATTTGAAAAAGTCAGGTGTAAATCTTAATACATCTGCTTTATATGTTAATG 1532

Db 130650 ATAGTCTCTGCT-GAAAAAGTCTGGTGTGATTTCTAATAGGCGCTGCCCTTATATGTTACTTG 130708
QY 1533 GTCTTTTTCCTTGGCATCTTTTAAATATCTTTCTTGTCTATATCTTTTAGTATTTGA 1592
Db 130709 ACCTTTTTTCCT-TTATTTGCTTTTAAATATCTATCTTTTAT-TGGTGTTTTGTGTTCTGA 130766
QY 1593 TTATTTATGCACTGTGGGAGTCTTTTTCGGTCCCAATCTATTTGGTGTGTTTGTATGCTTT 1652
Db 130767 TTAATATGTGTAGGAGGAATTTCTTTCTAGTCCAGTCTATTTGGAGTCTGTAGGCTTT 130826
QY 1653 CTGTGACCTGTAGGAGATCTTTCTCAAGGTAGGAAATTTTCTTTTGGTGTGTTTCT 1712
Db 130827 CTTGTATGTTTCATGGGCATCTCTTTCTTTAGGATGGGAAGTTTCTTCTATAAATTTCT 130886
QY 1713 TGAATAATTTTCCCTGCTTTTTCACCTG-----CCTTCTTCCCTTCTCTCT 1758
Db 130887 TGAAGATATTTGCTGGCCCTTTTAAAGTTGAAATCTTTCTTCTGCTCTACTCTATATAT 130946
QY 1759 ATTCTTTGGTGTGTTTGGATAGTCTCTGCTTCCCTGGAGTGTGTTTATGCTGCTGATTTT 1818
Db 130947 GTAGGTTTGTCTTCTCAITGTCTCTAAATTTCTTAGATGTTTTCAGTCCAGATCTTTT 131006
QY 1819 TAGACTTAAATTTTCTTTGACCAAGGTATCCATTTCTTCTATCTGTGTCTTCACTGCCCTG 1878
Db 131007 TGCATTTGCAATTTCTTTGATGTTGTGCGGATGTCTCTATGTAATCTTCTGACCTG 131066
QY 1879 AGATTCCTCTTCTATCTCTTGTATCTGTGAGTGTGCTGTGAGGCTTGTCTGAGGTTCTCTG 1936
Db 131067 AGATTCCTCTTCCATCTCTTGTATTTCTGTGTAATGCTCGCATCTATGGTTCTAGATT 131126
QY 1937 TGGGTTCTTAATTTTCTTCAATTTCCAGATTTCTTCAAGTTGGGTTGTTGTTTATTAATCT 1996
Db 131127 TTTTCCCTAGGTTTCTATCTCCAGGTTGCTCACTTTGGGTTTCTTTATTTGTGTCT 131186
QY 1997 ATTTCCACTTTTCCAGTCTCTGAAATGTTTACTCAITTT---TCCTCCAGATTTTACATTTT 2054
Db 131187 ACTTCACTTTTAGGCTTGTATGATTTTGTTCATTTCCACCACCTGTTTGTGTGTTT 131246
QY 2055 CATAGTCTTCTTTAATGGAATTTATTCATTTCTCTTTCA 2092
Db 131247 TCCTGTTTCTCTAAGGATTTGTACTGTACCTCTTTA 131284

RESULT 36

US-10-175-523-75
; Sequence 75, Application US/10175523
; Publication No. US20030096264A1
; GENERAL INFORMATION:
; APPLICANT: Brockman, David
; APPLICANT: Evans, David
; APPLICANT: Hook, Derek
; APPLICANT: Klimczak, Leszek
; APPLICANT: Laeng, Pascal
; APPLICANT: Palfreyman, Michael
; APPLICANT: Rajan, Prithi
; TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)
; FILE REFERENCE: 3235/10795-US3
; CURRENT APPLICATION NUMBER: US/10/175,523
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 60/299,151
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/317,828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/325,150
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/333,047
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 60/349,936
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/361,834
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1